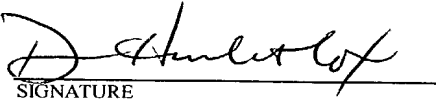


JC07 Rec'd PCT/PTO 04 DEC 2001

FORM PTO-1390 U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER <b>PF-0709 USN</b>
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371		U.S. APPLICATION NO. <b>10/009328</b> TO BE ASSIGNED
INTERNATIONAL APPLICATION NO PCT/US00/16668	INTERNATIONAL FILING DATE 16 June 2000	PRIORITY DATE CLAIMED 17 June 1999
TITLE OF INVENTION <b>HUMAN TRANSPORT PROTEINS</b>		
APPLICANT(S) FOR DO/EO/US <b>INCYTE GENOMICS, INC.; LAL, Preeti; YANG, Junming; YUE, Henry; HILLMAN, Jennifer L.; TANG, Y. Tom; BANDMAN, Olga; BURFORD, Neil; BAUGHN, Mariah R.; AZIMZAI, Yalda; LU, Dyung Aina M.; AU-YOUNG, Janice; PATTERSON, Chandra</b>		
<p>Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:</p> <ol style="list-style-type: none"> <li><input checked="" type="checkbox"/> This is the <b>FIRST</b> submission of items concerning a filing under 35 U.S.C. 371.</li> <li><input type="checkbox"/> This is a <b>SECOND</b> or <b>SUBSEQUENT</b> submission of items concerning a filing under 35 U.S.C. 371.</li> <li><input type="checkbox"/> This is an express request to promptly begin national examination procedures (35 U.S.C. 371 (f)).</li> <li><input type="checkbox"/> The US has been elected by the expiration of 19 months from the priority date (PCT Article 31).</li> <li><input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2))             <ol style="list-style-type: none"> <li><input type="checkbox"/> is attached hereto (required only if not communicated by the International Bureau)</li> <li><input type="checkbox"/> has been communicated by the International Bureau.</li> <li><input checked="" type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US).</li> </ol> </li> <li><input type="checkbox"/> An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).</li> <li><input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))             <ol style="list-style-type: none"> <li><input type="checkbox"/> are attached hereto (required only if not communicated by the International Bureau).</li> <li><input type="checkbox"/> have been communicated by the International Bureau.</li> <li><input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.</li> <li><input checked="" type="checkbox"/> have not been made and will not be made.</li> </ol> </li> <li><input type="checkbox"/> An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).</li> <li><input checked="" type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).</li> <li><input type="checkbox"/> An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).</li> </ol>		
<p><b>Items 11 to 16 below concern document(s) or information included:</b></p> <ol style="list-style-type: none"> <li><input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.</li> <li><input checked="" type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.27 and 3.31 is included.</li> <li><input checked="" type="checkbox"/> A <b>FIRST</b> preliminary amendment, as follows: Cancel in this application original claims 12, 14, 18, 20, 21, 23, 24, 25, 28 - 230 before calculating the filing fee, without prejudice or disclaimer. Applicants submit that these claims were included in the application as filed in the interest of providing notice to the public of certain specific subject matter intended to be claimed, and are being canceled at this time in the interest of reducing filing costs. Applicants expressly state that these claims are not being canceled for reasons related to patentability, and are in fact fully supported by the specification as filed. Applicants expressly reserve the right to reinstate these claims or to add other claims during prosecution of this application or a continuation or divisional application. Applicants expressly do not disclaim the subject matter of any invention disclosed herein which is not set forth in the instantly filed claims.             <ol style="list-style-type: none"> <li><input type="checkbox"/> A <b>SECOND</b> or <b>SUBSEQUENT</b> preliminary amendment.</li> </ol> </li> <li><input type="checkbox"/> A substitute specification.</li> <li><input type="checkbox"/> A change of power of attorney and/or address letter.</li> <li><input checked="" type="checkbox"/> Other items or information:             <ol style="list-style-type: none"> <li>1) Transmittal Letter (2 pp, in duplicate)</li> <li>2) Return Postcard</li> <li>3) Express Mail Label No.: <b>EL 697 344 232 US</b></li> <li>4) Article 34 Amendment</li> </ol> </li> </ol>		

JC07 Rec'd PCT/PTO 04 DEC 2001

U.S. APPLICATION NO. (if known) see 37 CFR 1.5 <b>TO BE ASSIGNED 10/009328</b>		INTERNATIONAL APPLICATION NO. PCT/US00/16668		ATTORNEY'S DOCKET NUMBER PF-0709 USN	
17 <input checked="" type="checkbox"/> The following fees are submitted: <b>BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5)):</b> Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO.....\$1000.00 <input type="checkbox"/> International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO..\$860.00 International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO.....\$710.00 <input checked="" type="checkbox"/> <b>International preliminary examination fee paid to USPTO (37 CFR 1.482)</b> <b>but all claims did not satisfy provisions of PCT Article 33(1)-(4)).....\$710.00</b> <input type="checkbox"/> International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4)).....\$100.00					
ENTER APPROPRIATE BASIC FEE AMOUNT =				\$710.00	
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				\$	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total Claims	20 =	0	X \$ 18.00	\$	
Independent Claims	3 =	1	X \$ 80.00	\$ 80.00	
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ \$270.00	\$	
TOTAL OF ABOVE CALCULATIONS =				\$790.00	
<input type="checkbox"/> Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2.				\$	
SUBTOTAL =				\$790.00	
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(f)).				\$	
TOTAL NATIONAL FEE =				\$790.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by the appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property				\$	
TOTAL FEES ENCLOSED =				\$790.00	
				Amount to be Refunded	\$
				Charged	\$
a. <input type="checkbox"/> A check in the amount of \$_____ to cover the above fees is enclosed. b. <input checked="" type="checkbox"/> Please charge my Deposit Account No. <u>09-0108</u> in the amount of <u>\$790.00</u> to cover the above fees. c. <input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. <u>09-0108</u> . A duplicate copy of this sheet is enclosed.					
<b>NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must          be filed and granted to restore the application to pending status.</b>					
SEND ALL CORRESPONDENCE TO:  INCYTE GENOMICS, INC. 3160 Porter Drive Palo Alto, CA 94304					
 SIGNATURE					
NAME: Diana Hamlet-Cox					
REGISTRATION NUMBER: 33,302					
DATE: <u>4</u> December 2001					

"Express Mail" mailing label number EL 697 344 232 US. I hereby certify that this document and referenced attachments are being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR § 1.10, addressed to: Commissioner for Patents, Box Patent Application, P.O. Box 2327, Arlington, VA 22202 on 4 December 2001.

By: Nancy Ramos Printed: NANCY RAMOS

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

In re Application of: Preeti Lal, Junming Yang, Henry Yue, Jennifer L. Hillman, Y. Tom Tang, Olga Bandman, Neil Burford, Mariah R. Baughn, Yalda Azimzai, Dyung Aina M. Lu, Janice Au-Young & Chandra Patterson

Title: HUMAN TRANSPORT PROTEINS

PCT Serial No.: PCT/US00/16668

International Filing Date: 16 June 2000

Examiner: To Be Assigned

Group Art Unit: To Be Assigned

---

Assistant Commissioner for Patents

**Box Patent Application**

Washington, D.C. 20231

**SUBMISSION UNDER 37 CFR § 1.821-1.825 SEQUENCE LISTING**

Sir:

In accordance with the requirements of 37 CFR § 1.821-1.825, Applicants hereby submit one (1) diskette(s) containing the computer-readable information for the Sequence Listing of the above-identified application. The content of the Sequence Listing paper copy is identical to the computer-readable copy filed with the US Receiving Office. The USPTO is authorized to add whatever is necessary to update the CRF with the current application information.

Respectfully submitted,

INCYTE GENOMICS, INC.

Diana Hamlet-Cox

Diana Hamlet-Cox

Reg. No. 33,302

Direct Dial Telephone: (650) 845-4639

Date: 4 December 2001

3160 Porter Drive  
Palo Alto, California 94304  
Phone: (650) 855-0555  
Fax: (650) 845-4166

## TECHNICAL FIELD

10 BACKGROUND OF THE INVENTION

Cells and organelles require transport molecules to import and export essential nutrients and metal ions including  $K^+$ ,  $NH_4^+$ ,  $P_i$ ,  $SO_4^{2-}$ , sugars, and vitamins, as well as various metabolic waste products. Transport proteins also play roles in antibiotic resistance, toxin secretion, ion balance, synaptic neurotransmission, kidney function, intestinal absorption, tumor growth, and other diverse cell functions (Griffith, J. and C. Sansom (1998) The Transporter Facts Book, Academic Press, San Diego CA, pp. 3-29). Transport can occur by a passive concentration-dependent mechanism, or can be linked to an energy source such as ATP hydrolysis or an ion gradient. Proteins that function in transport include carrier proteins, which bind to a specific solute and undergo a conformational change that transfers the bound solute across the membrane, and channel proteins, which form hydrophilic pores that allow specific solutes to diffuse through the membrane down an electrochemical solute gradient.

35 Active transport is tightly coupled to a source of metabolic energy, such as ATP hydrolysis or an electrochemically favorable ion gradient. Passive transport involves the movement of a solute down



WO 00/78953

PCT/US00/16668

its electrochemical gradient. Transport proteins can be further classified as either carrier proteins or channel proteins. Carrier proteins, which can function in active or passive transport, bind to a specific solute to be transported and undergo a conformational change which transfers the bound solute across the membrane. Channel proteins, which only function in passive transport, form hydrophilic pores across the membrane. When the pores open, specific solutes, such as inorganic ions, pass through the membrane and down the electrochemical gradient of the solute. Examples include facilitative transporters, the secondary active symporters and antiporters driven by ion gradients, and active ATP binding cassette transporters involved in multiple-drug resistance and targeting of antigenic peptides to MHC Class I molecules. Transported substrates range from nutrients and ions to a broad variety of drugs, peptides and proteins.

Information on the action of ARL-6 (ADP-ribosylation like factor), an endoplasmic reticulum transmembrane protein, can be found in Greenfield, J.J. and S. High (1999; J. Cell Sci. 112:1477-1486). Information on reduced folate carrier transporter proteins can be found in Dixon, K.H. et al. (1994; J. Biol. Chem. 269:17-20) and Moscow, J.A. et al. (1995; Cancer Res. 55:5983-5987).

Carrier proteins which transport a single solute from one side of the membrane to the other are called uniporters. In contrast, coupled transporters link the transfer of one solute with simultaneous or sequential transfer of a second solute, either in the same direction (symport) or in the opposite direction (antiport). For example, intestinal and kidney epithelia contain a variety of symporter systems driven by the sodium gradient that exists across the plasma membrane. Sodium moves into the cell down its electrochemical gradient and brings the solute into the cell with it. The sodium gradient that provides the driving force for solute uptake is maintained by the ubiquitous  $\text{Na}^+/\text{K}^+$  ATPase. Sodium-coupled transporters include the mammalian glucose transporter (SGLT1), iodide transporter (NIS), and multivitamin transporter (SMVT). All three transporters have twelve putative transmembrane segments, extracellular glycosylation sites, and cytoplasmically-oriented N- and C-termini.

Mitochondrial carrier proteins are transmembrane-spanning proteins which transport ions and charged metabolites between the cytosol and the mitochondrial matrix. Examples include the ADP, ATP carrier protein; the 2-oxoglutarate/malate carrier; the phosphate carrier protein; the brown fat uncoupling protein which transports protons from the cytosol into the matrix; the pyruvate carrier; the dicarboxylate carrier which transports malate, succinate, fumarate, and phosphate; the tricarboxylate carrier which transports citrate and malate; and the Grave's disease carrier protein, a protein recognized by IgG in patients with active Grave's disease, an autoimmune disorder resulting in hyperthyroidism (Stryer, L. (1995) Biochemistry, W.H. Freeman and Company, New York NY, p. 551; PROSITE PDOC00189 Mitochondrial energy transfer proteins signature; Online Mendelian Inheritance in Man (OMIM) \*275000 Graves Disease).

This class of transporters also includes the mitochondrial uncoupling proteins, which create

proton leaks across the inner mitochondrial membrane, thus uncoupling oxidative phosphorylation from ATP synthesis. The result is energy dissipation in the form of heat. Mitochondrial uncoupling proteins have been implicated as modulators of thermoregulation and metabolic rate, and have been proposed as potential targets for drugs against metabolic diseases such as obesity (Ricquier, D. et al. (1999) J. Int. Med. 245:637-642).

A number of metal ions such as iron, zinc, copper, cobalt, manganese, molybdenum, selenium, nickel, and chromium are important as cofactors for a number of enzymes. For example, zinc is required for the function of enzymes such as the extracellular matrix metalloproteinases, and zinc ions stabilize several motifs commonly found in transcription factors, including zinc fingers, zinc clusters, and LIM domains. Zinc and other metal ions must be provided in the diet, and are absorbed by transporters in the gastrointestinal tract. Plasma proteins transport the metal ions to the liver and other target organs, where specific transporters move the ions into cells and cellular organelles as needed. Imbalances in metal ion metabolism have been associated with a number of disease states (Danks, D.M. (1986) J. Med. Genet. 23:99-106).

The largest and most diverse family of transport proteins known are the ATP-binding cassette (ABC) transporters. As a family, ABC transporters can transport substances that differ markedly in chemical structure and size, ranging from small molecules such as ions, sugars, amino acids, peptides, and phospholipids, to lipopeptides, large proteins, and complex hydrophobic drugs. ABC proteins consist of four modules: two nucleotide-binding domains (NBD), which hydrolyze ATP to supply the energy required for transport, and two membrane-spanning domains (MSD), each containing six putative transmembrane segments. These four modules may be encoded by a single gene, as is the case for the cystic fibrosis transmembrane regulator (CFTR), or by separate genes. When encoded by separate genes, each gene product contains a single NBD and MSD. These "half-molecules" form homo- and heterodimers, such as Tap1 and Tap2, the endoplasmic reticulum-based major histocompatibility (MHC) peptide transport system. Several genetic diseases are attributed to defects in ABC transporters, such as the following diseases and their corresponding proteins: cystic fibrosis (CFTR, an ion channel), adrenoleukodystrophy (adrenoleukodystrophy protein, ALDP), Zellweger syndrome (peroxisomal membrane protein-70, PMP70), and hyperinsulinemic hypoglycemia (sulfonylurea receptor, SUR). Overexpression of the multidrug resistance (MDR) protein, another ABC transporter, in human cancer cells makes the cells resistant to a variety of cytotoxic drugs used in chemotherapy (Taglicht, D. and S. Michaelis (1998) Methods Enzymol. 292:131-163).

The nuclear pore complex (NPC) is a large multiprotein complex spanning the nuclear envelope which mediates the transport of proteins and RNA molecules between the nucleus and the cytoplasm, thus contributing to the regulation of gene expression. The NPC allows passive diffusion of ions, small molecules, and macromolecules under about 60kD, while larger macromolecules are transported by facilitated, energy-dependent pathways. Nuclear localization signals (NLS), consisting

of short stretches of amino acids enriched in basic residues, are found on proteins that are targeted to the nucleus, such as the glucocorticoid receptor. The NLS is recognized by the NLS receptor, importin, which then interacts with the monomeric GTP-binding protein Ran. This NLS protein/receptor/Ran complex navigates the nuclear pore with the help of the homodimeric protein nuclear transport factor 2 (NTF2) (Nakielnny, S. and G. Dreyfuss (1997) *Curr. Opin. Cell Biol.* 9:420-429; Gorlich, D. (1997) *Curr. Opin. Cell Biol.* 9:412-419). Four O-linked glycoproteins, p62, p58, p54, and p45, exist as a stable "p62 complex" that forms a ring localized on both nucleoplasmic and cytoplasmic surfaces of the NPC. The p62, p58, and p54 proteins all interact directly with the cytosolic transport factors p97 and NTF2, suggesting that the p62 complex is an important ligand binding site near the central gated channel of the NPC (Hu, T. et al. (1996) *J. Cell Biol.* 134:589-601).

Transport can also occur through intercellular bridges which connect the cytoplasms of sister cells, for example in the male and female germline of species ranging from fruit flies to humans. These bridges allow passage of cytoplasmic materials between cells during development. Intercellular bridges have also been found to connect somatic cells. The nurse cells and oocyte of a *Drosophila* egg chamber, which are derived from a single precursor cell through four rounds of mitosis, are connected to each other through intercellular bridges called ring canals. The cells do not completely separate after mitosis; the mitotic cleavage furrows are transformed into ring canals by the addition of an actin cytoskeleton lining the tunnels between the cells. The *Drosophila* kelch protein functions in organizing actin in the ring canal. Mutations in kelch cause female sterility in *Drosophila*. Kelch contains four protein domains: the NTR domain at the N-terminus, the BTB or POZ domain, the IVR or intervening region; and the kelch repeat domain, which contains six 50-amino acid kelch repeats. The BTB or POZ domain, a 120-amino acid motif that is also found in several zinc-finger containing transcription factors, may be important in dimerization of kelch. Kelch repeats are found in other proteins as well and may be important for actin binding (Robinson, D.N. and L. Cooley (1997) *J. Cell Biol.* 138:799-810; Cooley, L. (1998) *Cell* 93:913-915).

### Ion Channels

The electrical potential of a cell is generated and maintained by controlling the movement of ions across the plasma membrane. The movement of ions requires ion channels, which form an ion-selective pore within the membrane. Ion channels share common structural and mechanistic themes. The channel consists of four or five subunits or protein monomers that are arranged like a barrel in the plasma membrane. Each subunit typically consists of six potential transmembrane segments (S1, S2, S3, S4, S5, and S6). The center of the barrel forms a pore lined by  $\alpha$ -helices or  $\beta$ -strands. The side chains of the amino acid residues comprising the  $\alpha$ -helices or  $\beta$ -strands establish the charge (cation or anion) selectivity of the channel. The degree of selectivity, or what specific ions are allowed to pass through the channel, depends on the diameter of the narrowest part of the pore. There

are two basic types of ion channels, ion transporters and gated ion channels. Ion transporters utilize the energy obtained from ATP hydrolysis to actively transport an ion against the ion's concentration gradient. Gated ion channels allow passive flow of an ion down the ion's electrochemical gradient under restricted conditions. Together, these types of ion channels generate, maintain, and utilize an electrochemical gradient that is used in 1) electrical impulse conduction down the axon of a nerve cell, 2) transport of molecules into cells against concentration gradients, 3) initiation of muscle contraction, and 4) endocrine cell secretion.

Transmembrane ATPases are divided into three families. The phosphorylated (P) class ion transporters, including Na<sup>+</sup>-K<sup>+</sup> ATPase, Ca<sup>2+</sup> ATPase, H<sup>+</sup> ATPase, and Cu<sup>++</sup> ATPase, are activated by a phosphorylation event. P-class ion transporters are responsible for maintaining resting potential distributions such that cytosolic concentrations of Na<sup>+</sup> and Ca<sup>2+</sup> are low and cytosolic concentration of K<sup>+</sup> is high. The vacuolar (V) class of ion transporters include H<sup>+</sup> pumps on intracellular organelles, such as lysosomes and Golgi. V-class ion transporters are responsible for generating the low pH within the lumen of these organelles that is required for function. The coupling factor (F) class consists of H<sup>+</sup> pumps in the mitochondria. F-class ion transporters utilize a proton gradient to generate ATP from ADP and inorganic phosphate (P<sub>i</sub>).

Cu<sup>++</sup> ATPases export copper from cells (PROSITE PDOC00139 E1-E2 ATPases phosphorylation site). Mutations in one Cu<sup>++</sup> ATPase cause Wilson disease, in which toxic amounts of copper accumulate in a number of organs, particularly the liver and brain (Tanzi, R.E. et al. (1993) Nat. Genet. 5:344-350). Mutations in another Cu<sup>++</sup> ATPase cause Menkes disease and occipital horn syndrome. Menkes disease mutations block export of copper from the gastrointestinal tract, leading to skeletal abnormalities, severe mental retardation, neurologic degeneration, and mortality in early childhood (Harrison, M.D. and C.T. Dameron (1999) J. Biochem. Mol. Toxicol. 13:93-106). Occipital horn syndrome mutations cause connective tissue defects (Harrison, supra; Levinson, B. et al. (1996) Hum. Mol. Genet. 5:1737-1742).

The coupling factor (F) class of ion transporters consists of H<sup>+</sup> pumps in mitochondria, chloroplasts, and bacteria. For example, the F<sub>0</sub>F<sub>1</sub> ATPase utilizes a proton gradient across the inner mitochondrial membrane to generate ATP from ADP and inorganic phosphate (P<sub>i</sub>). The F<sub>0</sub>F<sub>1</sub> ATPase is composed of the F<sub>0</sub> complex, which is the transmembrane channel through which protons flow, and the F<sub>1</sub> complex, where ATP synthesis activity resides. F<sub>0</sub> has three subunits, A (also known as protein 6), B, and C (Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York NY, pp. 752-756; PROSITE PDOC00420 ATP synthase a subunit signature).

Voltage-gated Ca<sup>2+</sup> channels are involved in presynaptic neurotransmitter release, and heart and skeletal muscle contraction. The voltage-gated Ca<sup>2+</sup> channels from skeletal muscle (L-type) and brain (N-type) have been purified and, though their functions differ dramatically, they have similar subunit compositions. The channels are composed of three subunits. The α<sub>1</sub> subunit forms the

membrane pore and voltage sensor, while the  $\alpha_2\delta$  and  $\beta$  subunits modulate the voltage-dependence, gating properties, and the current amplitude of the channel. These subunits are encoded by at least six  $\alpha_1$ , one  $\alpha_2\delta$ , and four  $\beta$  genes. A fourth subunit,  $\gamma$ , has been identified in skeletal muscle (Walker, D. et al. (1998) J. Biol. Chem. 273:2361-2367; and Jay, S.D. et al. (1990) Science 248:490-492). The human  $\beta_4$  subunit is homologous to the mouse epilepsy gene lethargic, and is a candidate for involvement in neurological disorders including ataxia and absence epilepsy (Escayg, A. et al. (1998) Genomics 50:14-22).

Ligand-gated channels open their pores when an extracellular or intracellular mediator binds to the channel. Neurotransmitter-gated channels are channels that open when a neurotransmitter binds to their extracellular domain. These channels exist in the postsynaptic membrane of nerve or muscle cells. There are two types of neurotransmitter-gated channels. Sodium channels open in response to excitatory neurotransmitters, such as acetylcholine, glutamate, and serotonin. This opening causes an influx of  $\text{Na}^+$  and produces the initial localized depolarization that activates the voltage-gated channels and starts the action potential. Chloride channels open in response to inhibitory neurotransmitters, such as  $\gamma$ -aminobutyric acid (GABA) and glycine, leading to hyperpolarization of the membrane and the subsequent generation of an action potential.

Ion channels are expressed in a number of tissues where they are implicated in a variety of processes. CNG channels, while abundantly expressed in photoreceptor and olfactory sensory cells, are also found in kidney, lung, pineal, retinal ganglion cells, testis, aorta, and brain. Calcium-activated  $\text{K}^+$  channels may be responsible for the vasodilatory effects of bradykinin in the kidney and for shunting excess  $\text{K}^+$  from brain capillary endothelial cells into the blood. They are also implicated in repolarizing granulocytes after agonist-stimulated depolarization (Ishi, T.M. et al. (1997) Proc. Natl. Acad. Sci. USA 94:11651-11656). Another transmembrane protein, the leukotrine B4 receptor (BLT) appears to be involved in inflammation responses and host cell defense against infection. BLT also functions as an HIV coreceptor (Izumi, T. et al. (1997) Nature 387:620-624; Martin, V. et al. (1999) J. Biol. Chem. 274:8597-8603).

Ion channels have been the target for many drug therapies. Neurotransmitter-gated channels have been targeted in therapies for treatment of insomnia, anxiety, depression, and schizophrenia. Voltage-gated channels have been targeted in therapies for arrhythmia, ischemic stroke, head trauma, and neurodegenerative disease (Taylor, C.P. and L.S. Narasimhan (1997) Adv. Pharmacol. 39:47-98).

$\text{K}^+$  channels are located in all cell types, and may be regulated by voltage, ATP concentration, or second messengers such as  $\text{Ca}^{++}$  and cAMP. In non-excitabile tissue,  $\text{K}^+$  channels are involved in protein synthesis, control of endocrine secretions, and the maintenance of osmotic equilibrium across membranes. In neurons and other excitable cells, in addition to regulating action potentials and repolarizing membranes,  $\text{K}^+$  channels are responsible for setting resting membrane potential. The cytosol contains non-diffusible anions and, to balance this net negative charge, the cell

WO 00/78953

PCT/US00/16668

contains a Na<sup>+</sup>-K<sup>+</sup> pump and ion channels that provide the redistribution of Na<sup>+</sup>, K<sup>+</sup>, and Cl<sup>-</sup>. The pump actively transports Na<sup>+</sup> out of the cell and K<sup>+</sup> into the cell in a 3:2 ratio. Ion channels in the plasma membrane allow K<sup>+</sup> and Cl<sup>-</sup> to flow by passive diffusion. Because of the high negative charge within the cytosol, Cl<sup>-</sup> flows out of the cell. The flow of K<sup>+</sup> is balanced by an electromotive force  
5 pulling K<sup>+</sup> into the cell, and a K<sup>+</sup> concentration gradient pushing K<sup>+</sup> out of the cell. Thus, the resting membrane potential is primarily regulated by K<sup>+</sup> flow (Salkoff, L. and T. Jegla (1995) *Neuron* 15:489-492). Information on NY-REN-45, a K<sup>+</sup> channel integral membrane protein, can be found in Scanlan, M.J. et al. (1998; *Int. J. Cancer* 76:652-658). The emopamil-binding protein (EBP) shares structural features with both pro- and eukaryotic drug transport proteins (Hanner, M. et al. (1995) *J. Biol. Chem.* 270:7551-7557). The Na<sup>+</sup> channel, transmembrane protein myelin protein zero (MPZ)  
10 may be responsible for some sporadic cases of Dejerine-Scottas disease (hereditary motor and sensory neuropathy type III) (Hayasaka, K. et al. (1993) *Nat. Genet.* 5:266-268).

K<sup>+</sup> pore-forming subunits generally have six transmembrane-spanning domains with a short region between the fifth and sixth transmembrane regions that senses membrane potential; and the  
15 amino and carboxy termini are located intracellularly. In mammalian heart, the duration of ventricular action potential is controlled by a K<sup>+</sup> current. Thus, the K<sup>+</sup> channel is central to the control of heart rate and rhythm. K<sup>+</sup> channel dysfunctions are associated with a number of renal diseases including hypertension, hypokalemia, and the associated Bartter's syndrome and Getelman's syndrome, as well as neurological disorders including epilepsy. K<sup>+</sup> channels have been implicated in  
20 Alzheimer's disease by observations that a significant component of senile plaques, beta amyloid or A beta, also blocks voltage-gated potassium channels in hippocampal neurons (Antes, L.M. et al. (1998) *Seminars Nephrol.* 18:31-45; Stoffel, M. and L.Y. Jan (1998) *Nat. Genet.* 18:6-8; Madeja, M. et al. (1997) *Eur. J. Neurosci.* 9:390-395; Good, T.A. et al. (1996) *Biophys. J.* 70:296-304).

Gated ion channels control ion flow by regulating the opening and closing of pores. These  
25 channels are categorized according to the manner of regulating the gating function. Mechanically-gated channels open pores in response to mechanical stress, voltage-gated channels open pores in response to changes in membrane potential, and ligand-gated channels open pores in the presence of a specific ion, nucleotide, or neurotransmitter.

Voltage-gated Na<sup>+</sup> channels are responsible for electrical excitability of neurons, skeletal  
30 muscle, heart, and neuroendocrine tissues. For example, the sequential opening and closing of voltage-gated Na<sup>+</sup> channels results in the propagation of action potentials down neuronal axons. Na<sup>+</sup> channels isolated from rat brain tissue are heterotrimeric complexes composed of a 260 kDa pore forming  $\alpha$  subunit that associates with two smaller auxiliary subunits,  $\beta$ 1 and  $\beta$ 2. The  $\beta$ 2 subunit is an integral membrane glycoprotein that contains an extracellular Ig domain, and its association with  $\alpha$   
35 and  $\beta$ 1 subunits correlates with increased function of the channel, a change in the channel's gating properties, as well as an increase in whole cell capacitance (Isom, L.L. et al. (1995) *Cell* 83:433-442).

### Integral Membrane Proteins

The majority of known integral membrane proteins are transmembrane proteins (TM) which are characterized by an extracellular, a transmembrane, and an intracellular domain. TM domains are typically comprised of 15 to 25 hydrophobic amino acids which are predicted to adopt an  $\alpha$ -helical conformation. TM proteins are classified as bitopic (Types I and II) and polytopic (Types III and IV) (Singer, S.J. (1990) Annu. Rev. Cell Biol. 6:247-96). Bitopic proteins span the membrane once while polytopic proteins contain multiple membrane-spanning segments. TM proteins that act as cell-surface receptor proteins involved in signal transduction include growth and differentiation factor receptors, and receptor-interacting proteins such as Drosophila pecanex and frizzled proteins, LIV-1 protein, NF2 protein, and GNS1/SUR4 eukaryotic integral membrane proteins. TM proteins also act as transporters of ions or metabolites, such as gap junction channels (connexins) and ion channels, and as cell anchoring proteins, such as lectins, integrins, and fibronectins. TM proteins act as vesicle organelle-forming molecules, such as calveolins, or as cell recognition molecules, such as cluster of differentiation (CD) antigens, glycoproteins, and mucins. Information on connexin can be found in Kanter, H.L. et al. (1994; J. Mol. Cell. Cardiol. 26:861-868).

Many membrane proteins (MPs) contain amino acid sequence motifs that target these proteins to specific subcellular sites. Examples of these motifs include PDZ domains, KDEL, RGD, NGR, and GSL sequence motifs, von Willebrand factor A (vWFA) domains, and EGF-like domains. RGD, NGR, and GSL motif-containing peptides have been used as drug delivery agents in cancer treatments which target tumor vasculature (Arap, W. et al. (1998) Science, 279:377-380.) Furthermore, MPs may also contain amino acid sequence motifs, such as the carbohydrate recognition domain (CRD), also known as the C-type lectin domain, that mediate interactions with extracellular or intracellular molecules.

G-protein coupled receptors (GPCR) comprise a superfamily of integral membrane proteins which transduce extracellular signals. GPCRs include receptors for biogenic amines, lipid mediators of inflammation, peptide hormones, and sensory signal mediators. The structure of these highly-conserved receptors consists of seven hydrophobic transmembrane regions, an extracellular N-terminus, and a cytoplasmic C-terminus. Three extracellular loops alternate with three intracellular loops to link the seven transmembrane regions. The most conserved parts of these proteins are the transmembrane regions and the first two cytoplasmic loops. Cysteine disulfide bridges connect the second and third extracellular loops. A conserved, acidic-Arg-aromatic residue triplet present in the second cytoplasmic loop may interact with G proteins. A GPCR consensus pattern is characteristic of most proteins belonging to this superfamily (ExpASY PROSITE document PS00237; and Watson, S. and S. Arkinstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp 2-6). Mutations and changes in transcriptional activation of GPCR-encoding genes have been

associated with neurological disorders such as schizophrenia, Parkinson's disease, Alzheimer's disease, drug addiction, and feeding disorders.

Cytochromes are electron-transferring proteins that contain a heme prosthetic group, a porphyrin ring containing a tightly bound iron atom. Cytochromes act as oxidoreductases in such  
 5 diverse cellular processes as respiration, photosynthesis, fatty acid metabolism, and neurotransmitter biosynthesis. The heme iron atom serves as the actual electron carrier by changing from the ferric to the ferrous oxidation state when accepting an electron. Cytochromes accept electrons from one substrate such as NADH or ascorbate and donate them to other electron carriers such as other  
 10 cytochromes, ubiquinone, or semidehydroascorbic acid (Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York NY, pp. 759-770, 786-797; Sperling, P. et al. (1995) Eur. J. Biochem. 232:798-805; and Online Mendelian Inheritance in Man (OMIM) \*600019 Cytochrome b561, CYB561).

Cytochrome b5 is an electron donor in membrane-linked redox enzyme systems involved in lipid and drug metabolism. Cytochrome b5 has been found in Golgi, plasma, outer mitochondrial,  
 15 endoplasmic reticulum (ER), and microbody membranes. Conserved amino acids in cytochrome b5 include eight invariant amino acids at W34, H51, P52, G53, G54, G63, F70, and H74, and fifteen conserved amino acids at L24, I35, S36, V41, Y42, N43, T45, W47, A48, L58, D65, T67, L85, T87, and G88 (numbering based on the sunflower cytochrome b5/delta-6 desaturase fusion protein; GI 1040729, Sperling, supra). The invariant residues H51PGG are involved in heme-binding.  
 20 Cytochrome b5-like domains have also been found linked to other enzymes. For example, cytochrome b5-like domains are part of delta-9 fatty acid desaturases in yeast and Histoplasma capsulatum, nitrate reductase, sulfite reductase, flavocytochrome b2, Arabidopsis thaliana acyl lipid desaturase, and Borago officinalis (borage) and Helianthus annuus (sunflower) delta-6 desaturases (Sperling, supra; Sayanova, O. et al (1997) Proc. Natl. Acad. Sci. USA 94:4211-4216; and Mitchell,  
 25 A.G. and C.E. Martin (1997) J. Biol. Chem. 272:28281-28288).

Signal peptides are found on proteins that are targeted to the endoplasmic reticulum (ER). Signal peptides consist of stretches of amino acids enriched in hydrophobic residues. Signal peptides are usually found at the extreme N-terminus of the protein and are recognized by a cytosolic signal-recognition peptide (SRP). The SRP binds to the signal peptide and to an SRP receptor, an integral  
 30 membrane protein in the ER. Once bound to the SRP receptor, the newly formed protein containing the signal peptide is translocated across the ER membrane. Proteins containing signal peptides may end up inserted into the lipid bilayer, or they may end up in the lumen of an organelle or secreted from the cell.

### 35 Disease Correlation

The etiology of numerous human diseases and disorders can be attributed to defects in the



WO 00/78953

PCT/US00/16668

transport of molecules across membranes. Defects in the trafficking of membrane-bound transporters and ion channels are associated with several disorders, e.g. cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, von Gierke disease, and certain forms of diabetes mellitus. Single-gene defect diseases resulting in an inability to transport small molecules across  
5 membranes include, e.g., cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease (van't Hoff, W.G. (1996) Exp. Nephrol. 4:253-262; Talente, G.M. et al. (1994) Ann. Intern. Med. 120:218-226; and Chillon, M. et al. (1995) New Engl. J. Med. 332:1475-1480).

Cystinuria is an inherited disease that results from the inability to transport cystine, the disulfide-linked dimer of cysteine, from the urine into the blood. Accumulation of cystine in the  
10 urine leads to the formation of cystine stones in the kidneys.

Transthyretin (TTR), present in human plasma, binds to and transports the thyroid hormone thyroxine. Mutations in TTR result in the conversion of TTR to amyloid, an insoluble fibrillar structure. The resulting amyloid plaques have been shown to be the causative agent in the development of familial amyloid polyneuropathy and senile systemic amyloidosis (Miroy, G.J. et al.  
15 (1996) Proc. Natl. Acad. Sci. USA 93:15051-15056).

Stomatin, a 31-kDa erythrocyte integral membrane protein has been linked to the hereditary anemia stomatocytosis. This anemia is characterized by red blood cells that lack stomatin and leak Na<sup>+</sup> and K<sup>+</sup>. Thus, stomatin is presumed to play a role in the regulation of ion transport. Red blood cell ion transport defects are also linked to other disorders such as hypertension (Stewart, G.W.  
20 (1997) Int. J. Biochem. Cell Biol. 29:271-274).

The discovery of new human transport proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of transport, metabolic, neurological, cardiovascular, reproductive, and  
25 immune disorders, and cell proliferative disorders including cancer.

## SUMMARY OF THE INVENTION

The invention features purified polypeptides, human transport proteins, referred to collectively as "TPPT" and individually as "TPPT-1," "TPPT-2," "TPPT-3," "TPPT-4," "TPPT-5,"  
30 "TPPT-6," "TPPT-7," "TPPT-8," "TPPT-9," "TPPT-10," "TPPT-11," "TPPT-12," "TPPT-13," "TPPT-14," "TPPT-15," "TPPT-16," "TPPT-17," "TPPT-18," "TPPT-19," "TPPT-20," "TPPT-21," "TPPT-22," "TPPT-23," "TPPT-24," "TPPT-25," "TPPT-26," "TPPT-27," "TPPT-28," "TPPT-29," "TPPT-30," "TPPT-31," "TPPT-32," "TPPT-33," "TPPT-34," "TPPT-35," "TPPT-36," "TPPT-37," "TPPT-38," "TPPT-39," "TPPT-40," "TPPT-41," "TPPT-42," and "TPPT-43." In one aspect, the  
35 invention provides an isolated polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-

WO 00/78953

PCT/US00/16668

43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-43.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-43. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:44-86.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a

WO 00/78953

PCT/US00/16668

polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43.

The invention further provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:44-86, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a pharmaceutical composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and a pharmaceutically acceptable excipient. In one embodiment, the pharmaceutical composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional TPPT, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional TPPT, comprising administering to a patient in need of such treatment the pharmaceutical composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-43. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a pharmaceutical composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional TPPT, comprising administering to a patient in



Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

5 Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding TPPT were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

10

### DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing  
15 particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a  
20 reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be  
25 used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

### 30 DEFINITIONS

"TPPT" refers to the amino acid sequences of substantially purified TPPT obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of  
35 TPPT. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TPPT either by directly interacting with

TPPT or by acting on components of the biological pathway in which TPPT participates.

An "allelic variant" is an alternative form of the gene encoding TPPT. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or  
5 many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding TPPT include those sequences with deletions,  
10 insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as TPPT or a polypeptide with at least one functional characteristic of TPPT. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding TPPT, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding  
15 TPPT. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent TPPT. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of TPPT is retained. For example,  
20 negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

25 The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

30 "Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of TPPT. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small  
35 molecules, or any other compound or composition which modulates the activity of TPPT either by directly interacting with TPPT or by acting on components of the biological pathway in which TPPT

participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')<sub>2</sub>, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind TPPT polypeptides can be prepared using intact polypeptides or using fragments  
5 containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

10 The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen  
15 used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified  
20 sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or  
25 translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic TPPT, or of any oligopeptide thereof,  
30 to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

35 A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or



amino acid sequence. The composition may comprise a dry formulation or an aqueous solution.

Compositions comprising polynucleotide sequences encoding TPPT or fragments of TPPT may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be

5 deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (PE Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from  
10 one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least  
15 interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
20	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
25	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
30	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
35	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

40 Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

WO 00/78953

PCT/US00/16668

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

A "fragment" is a unique portion of TPPT or the polynucleotide encoding TPPT which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:44-86 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:44-86, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:44-86 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:44-86 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:44-86 and the region of SEQ ID NO:44-86 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-43 is encoded by a fragment of SEQ ID NO:44-86. A fragment of SEQ ID NO:1-43 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-43. For example, a fragment of SEQ ID NO:1-43 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-43. The precise length of a fragment of SEQ ID NO:1-43 and the region of SEQ ID NO:1-43 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended

purpose for the fragment.

A "full-length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full-length" polynucleotide sequence encodes a "full-length" polypeptide sequence.

5 "Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps  
10 in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of  
15 molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent  
20 similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at  
25 <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The  
30 "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

35 *Reward for match: 1*

*Penalty for mismatch: -2*

*Open Gap: 5 and Extension Gap: 2 penalties*

*Gap x drop-off: 50*

**Expect: 10**

**Word Size: 11**

5 *Filter: on*

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

15 Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases “percent identity” and “% identity,” as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and “diagonals saved”=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the “percent similarity” between aligned polypeptide sequence pairs.

30 Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the “BLAST 2 Sequences” tool Version 2.0.12 (Apr-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

**Matrix: BLOSUM62**

35 *Open Gap: 11 and Extension Gap: 1 penalties*

*Gap x drop-off: 50*

*Expect: 10*

*Word Size: 3*

*Filter: on*

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

“Human artificial chromosomes” (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term “humanized antibody” refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

“Hybridization” refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the “washing” step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating  $T_m$  and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.



polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an TPPT may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of TPPT.

"Probe" refers to nucleic acid sequences encoding TPPT, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2<sup>nd</sup> ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al., 1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs

WO 00/78953

PCT/US00/16668

can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, *supra*. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is







WO 00/78953

PCT/US00/16668

polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

## 10 THE INVENTION

The invention is based on the discovery of new human transport proteins (TPPT), the polynucleotides encoding TPPT, and the use of these compositions for the diagnosis, treatment, or prevention of transport, metabolic, neurological, cardiovascular, reproductive, and immune disorders, and cell proliferative disorders including cancer.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding TPPT. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each TPPT were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. In some cases, GenBank sequence identifiers are also shown in column 5. The Incyte clones and GenBank cDNA sequences, where indicated, in column 5 were used to assemble the consensus nucleotide sequence of each TPPT and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding TPPT. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:44-86

and to distinguish between SEQ ID NO:44-86 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express TPPT as a fraction of total tissues expressing TPPT. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing TPPT as a fraction of total tissues expressing TPPT. Column 5 lists the vectors used to subclone each cDNA library.

Of particular interest is the expression of SEQ ID NO:50 exclusively in cardiovascular tissue, the expression of SEQ ID NO:56 in nervous and gastrointestinal tissues, the expression of SEQ ID NO:57 in gastrointestinal tissues, and the expression of SEQ ID NO:66 in nervous system tissues. Of particular note is the tissue-specific expression of SEQ ID NO:75. Over 71% of the cDNA libraries expressing SEQ ID NO:75 are derived from lung tissue.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding TPPT were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:44 maps to chromosome 7 within the interval from 38.80 to 42.10 centiMorgans. SEQ ID NO:48 maps to chromosome X within the interval from 107.90 to 122.80 centiMorgans. SEQ ID NO:60 maps to chromosome 2 within the interval from 157.0 to 167.0 centiMorgans. SEQ ID NO:65 maps to chromosome 2 within the interval from 17.4 to 40.7 centiMorgans and to chromosome 5 within the interval from 61.1 to 69.6 centiMorgans. The interval on chromosome 5 from 61.1 to 69.6 centiMorgans also contains genes associated with Cockayne syndrome. SEQ ID NO:69 maps to chromosome 3 within the interval from 157.40 to 162.00 centiMorgans. SEQ ID NO:70 maps to chromosome 3 within the interval from 176.40 to 179.80 centiMorgans. SEQ ID NO:71 maps to chromosome 18 within the interval from the p-terminus to 52.30 centiMorgans. SEQ ID NO:73 maps to chromosome 17 within the interval from 75.70 to 84.20 centiMorgans, and to chromosome 2 within the interval from 204.70 to 209.30 centiMorgans. SEQ ID NO:76 maps to chromosome 20 within the interval from 79.00 to 94.40 centiMorgans. SEQ ID NO:80 maps to chromosome 18 within the interval from 1.60 to 6.20 centiMorgans, and to chromosome 11 within the interval from 117.90 to 126.00 centiMorgans. SEQ ID NO:83 maps to chromosome 17 within the interval from 67.60 to 69.30 centiMorgans, and from 83.8 centiMorgans to the q-terminus, and to chromosome 7 within the interval from 105.20 to 114.50 centiMorgans.

The invention also encompasses TPPT variants. A preferred TPPT variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the TPPT amino acid sequence, and which contains at least one functional or structural characteristic of TPPT.

The invention also encompasses polynucleotides which encode TPPT. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected

WO 00/78953

PCT/US00/16668

from the group consisting of SEQ ID NO:44-86, which encodes TPPT. The polynucleotide sequences of SEQ ID NO:44-86, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

5 The invention also encompasses a variant of a polynucleotide sequence encoding TPPT. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding TPPT. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:44-  
10 86 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:44-86. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of TPPT.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the  
15 genetic code, a multitude of polynucleotide sequences encoding TPPT, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the  
20 polynucleotide sequence of naturally occurring TPPT, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode TPPT and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring TPPT under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding TPPT or  
25 its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding TPPT and its derivatives without altering the encoded amino acid sequences  
30 include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode TPPT and TPPT derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems  
35 using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding TPPT or any fragment thereof.



Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, PE Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode TPPT may be cloned in recombinant DNA molecules that direct expression of TPPT, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express TPPT.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter TPPT-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or

WO 00/78953

PCT/US00/16668

improve the biological properties of TPPT, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding TPPT may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser.* 7:215-223; and Horn, T. et al. (1980) *Nucleic Acids Symp. Ser.* 7:225-232.) Alternatively, TPPT itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (PE Biosystems). Additionally, the amino acid sequence of TPPT, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active TPPT, the nucleotide sequences encoding TPPT or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding TPPT. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding TPPT. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding TPPT and its initiation codon and upstream regulatory sequences are inserted into



the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both  
 5 natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding TPPT and appropriate transcriptional and translational control  
 10 elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences  
 15 encoding TPPT. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or  
 20 animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184; Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945; Takamatsu, N. (1987) *EMBO J.* 6:307-311; Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984)  
 25 *Science* 224:838-843; Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences  
 30 to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al., (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344; Buller, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al. (1994) *Mol. Immunol.* 31(3):219-226; and Verma, I.M. and N. Somia (1997) *Nature* 389:239-242.) The invention is not limited by the host cell employed.

35 In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding TPPT. For example, routine cloning,

WO 00/78953

PCT/US00/16668

subcloning, and propagation of polynucleotide sequences encoding TPPT can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding TPPT into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of TPPT are needed, e.g. for the production of antibodies, vectors which direct high level expression of TPPT may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of TPPT. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, *supra*; and Scorer, *supra*.)

Plant systems may also be used for expression of TPPT. Transcription of sequences encoding TPPT may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, *supra*; Broglie, *supra*; and Winter, *supra*.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., *The McGraw Hill Yearbook of Science and Technology* (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding TPPT may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses TPPT in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet.

WO 00/78953

PCT/US00/16668

15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of TPPT in cell lines is preferred. For example, sequences encoding TPPT can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous  
 5 expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue  
 10 culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk<sup>-</sup>* and *ap<sup>r</sup>-* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or  
 15 herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which  
 20 alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system.  
 25 (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding TPPT is inserted within a marker gene sequence, transformed cells containing sequences encoding TPPT can be identified by the absence of marker gene function. Alternatively, a  
 30 marker gene can be placed in tandem with a sequence encoding TPPT under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding TPPT and that express TPPT may be identified by a variety of procedures known to those of skill in the art. These  
 35 procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or



American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding TPPT may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric TPPT protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of TPPT activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the TPPT encoding sequence and the heterologous protein sequence, so that TPPT may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled TPPT may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, <sup>35</sup>S-methionine.

TPPT of the present invention or fragments thereof may be used to screen for compounds that specifically bind to TPPT. At least one and up to a plurality of test compounds may be screened for specific binding to TPPT. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of TPPT, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which TPPT binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express TPPT, either as a secreted



to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding TPPT may also be manipulated *in vitro* in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding TPPT can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding TPPT is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress TPPT, e.g., by secreting TPPT in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

#### THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of TPPT and human transport proteins. In addition, the expression of TPPT is closely associated with neurological, cardiovascular, reproductive, gastrointestinal, and hematopoietic/immune tissues, and inflammation, cell proliferation, and cancer. Therefore, TPPT appears to play a role in transport, metabolic, neurological, cardiovascular, reproductive, and immune disorders, and cell proliferative disorders including cancer. In the treatment of disorders associated with increased TPPT expression or activity, it is desirable to decrease the expression or activity of TPPT. In the treatment of disorders associated with decreased TPPT expression or activity, it is desirable to increase the expression or activity of TPPT.

Therefore, in one embodiment, TPPT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TPPT. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer; cardiac disorders associated with transport, e.g., angina, bradyarrhythmia, tachyarrhythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline

WO 00/78953

PCT/US00/16668

myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis; neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia; and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease; a metabolic disorder such as Addison's disease, cerebrotendinous xanthomatosis, congenital adrenal hyperplasia, coumarin resistance, cystic fibrosis, diabetes, fatty hepatocirrhosis, fructose-1,6-diphosphatase deficiency, galactosemia, goiter, glucagonoma, glycogen storage diseases, hereditary fructose intolerance, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypoglycemia, hypothyroidism, hyperlipidemia, hyperlipemia, lipid myopathies, lipodystrophies, lysosomal storage diseases, mannosidosis, neuraminidase deficiency, obesity, pentosuria phenylketonuria, and pseudovitamin D-deficiency rickets; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins,



WO 00/78953

PCT/US00/16668

thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; an immune disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis,

pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, 5 fungal, parasitic, protozoal, and helminthic infections, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, 10 myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing TPPT or a fragment or derivative 15 thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TPPT including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified TPPT in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TPPT including, but not 20 limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of TPPT may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TPPT including, but not limited to, those listed above.

In a further embodiment, an antagonist of TPPT may be administered to a subject to treat or 25 prevent a disorder associated with increased expression or activity of TPPT. Examples of such disorders include, but are not limited to, those transport, metabolic, neurological, cardiovascular, reproductive, and immune disorders, and cell proliferative disorders including cancer described above. In one aspect, an antibody which specifically binds TPPT may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to 30 cells or tissues which express TPPT.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding TPPT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TPPT including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary 35 sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made

WO 00/78953

PCT/US00/16668

by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

- 5 An antagonist of TPPT may be produced using methods which are generally known in the art. In particular, purified TPPT may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind TPPT. Antibodies to TPPT may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and
- 10 fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

- For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with TPPT or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to
- 15 increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

- It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to TPPT
- 20 have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of TPPT amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

- 25 Monoclonal antibodies to TPPT may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and
- 30 Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

- In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda,
- 35 S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce TPPT-specific single

WO 00/78953

PCT/US00/16668

chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for TPPT may also be generated. For example, such fragments include, but are not limited to,  $F(ab')_2$  fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the  $F(ab')_2$  fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between TPPT and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering TPPT epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for TPPT. Affinity is expressed as an association constant,  $K_a$ , which is defined as the molar concentration of TPPT-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The  $K_a$  determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple TPPT epitopes, represents the average affinity, or avidity, of the antibodies for TPPT. The  $K_a$  determined for a preparation of monoclonal antibodies, which are monospecific for a particular TPPT epitope, represents a true measure of affinity. High-affinity antibody preparations with  $K_a$  ranging from about  $10^9$  to  $10^{12}$  L/mole are preferred for use in immunoassays in which the TPPT-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of TPPT, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to

WO 00/78953

PCT/US00/16668

determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of TPPT-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines  
 5 for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al., supra.)

In another embodiment of the invention, the polynucleotides encoding TPPT, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA,  
 10 RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding TPPT. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding TPPT. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense  
 15 sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral  
 20 vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.*  
 25 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding TPPT may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined  
 30 immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and Somia, N. (1997) *Nature* 389:239-242)), (ii)  
 35 express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated



In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to TPPT expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding TPPT under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4<sup>+</sup> T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding TPPT to cells which have one or more genetic abnormalities with respect to the expression of TPPT. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544; and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding TPPT to target cells which have one or more genetic abnormalities with respect to the expression of TPPT. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing TPPT to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has

been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding TPPT to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotech.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full-length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for TPPT into the alphavirus genome in place of the capsid-coding region results in the production of a large number of TPPT-coding RNAs and the synthesis of high levels of TPPT in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of TPPT into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have



been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

5        Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding TPPT.

10        Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of  
15        candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

      Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis.  
20        Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding TPPT. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

25        RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine,  
30        and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

      An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding TPPT. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not  
35        limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular

WO 00/78953

PCT/US00/16668

chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased TPPT expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding TPPT may be therapeutically useful, and in the treatment of disorders associated with decreased TPPT expression or activity, a compound which specifically promotes expression of the polynucleotide encoding TPPT may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding TPPT is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding TPPT are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding TPPT. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved

WO 00/78953

PCT/US00/16668

using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and  
5 monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest  
10 edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such pharmaceutical compositions may consist of TPPT, antibodies to TPPT, and mimetics, agonists, antagonists, or inhibitors of TPPT.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial,  
15 intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Pharmaceutical compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol  
20 delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration  
25 enhancers.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of pharmaceutical compositions may be prepared for direct intracellular  
30 delivery of macromolecules comprising TPPT or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, TPPT or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system  
35 (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell

WO 00/78953

PCT/US00/16668

culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

5 A therapeutically effective dose refers to that amount of active ingredient, for example TPPT or fragments thereof, antibodies of TPPT, and agonists, antagonists or inhibitors of TPPT, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population) or LD<sub>50</sub> (the dose  
10 lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD<sub>50</sub>/ED<sub>50</sub> ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED<sub>50</sub> with  
15 little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the  
20 severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

25 Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells,  
30 conditions, locations, etc.

## DIAGNOSTICS

In another embodiment, antibodies which specifically bind TPPT may be used for the diagnosis of disorders characterized by expression of TPPT, or in assays to monitor patients being treated with TPPT or agonists, antagonists, or inhibitors of TPPT. Antibodies useful for diagnostic  
35 purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for TPPT include methods which utilize the antibody and a label to detect TPPT in human body fluids

WO 00/78953

PCT/US00/16668

or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

5 A variety of protocols for measuring TPPT, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of TPPT expression. Normal or standard values for TPPT expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to TPPT under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of TPPT expressed in subject, 10 control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding TPPT may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and 15 quantify gene expression in biopsied tissues in which expression of TPPT may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of TPPT, and to monitor regulation of TPPT levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding TPPT or closely related molecules may be used to 20 identify nucleic acid sequences which encode TPPT. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding TPPT, allelic variants, or related sequences.

25 Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the TPPT encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:44-86 or from genomic sequences including promoters, enhancers, and introns of the TPPT gene.

Means for producing specific hybridization probes for DNAs encoding TPPT include the 30 cloning of polynucleotide sequences encoding TPPT or TPPT derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as <sup>32</sup>P or <sup>35</sup>S, or by enzymatic labels, 35 such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding TPPT may be used for the diagnosis of disorders

WO 00/78953

PCT/US00/16668

associated with expression of TPPT. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer; cardiac disorders associated with transport, e.g., angina, bradyarrhythmia, tachyarrhythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis; neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia; and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease; a metabolic disorder such as Addison's disease, cerebrotendinous xanthomatosis, congenital adrenal hyperplasia, coumarin resistance, cystic fibrosis, diabetes, fatty hepatocirrhosis, fructose-1,6-diphosphatase deficiency, galactosemia, goiter, glucagonoma, glycogen storage diseases, hereditary fructose intolerance, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypoglycemia, hypothyroidism, hyperlipidemia, hyperlipemia, lipid myopathies, lipodystrophies, lysosomal storage diseases, mannosidosis, neuraminidase deficiency, obesity, pentosuria phenylketonuria, and pseudovitamin D-deficiency rickets; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial

nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; an immune disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome,

WO 00/78953

PCT/US00/16668

allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding TPPT may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered TPPT expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding TPPT may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding TPPT may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding TPPT in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.



In order to provide a basis for the diagnosis of a disorder associated with expression of TPPT, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding TPPT, under conditions suitable for hybridization or amplification.

5 Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

10 Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

15 With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

20 Additional diagnostic uses for oligonucleotides designed from the sequences encoding TPPT may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding TPPT, or a fragment of a polynucleotide complementary to the polynucleotide encoding TPPT, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

25 In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding TPPT may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding TPPT are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are

WO 00/78953

PCT/US00/16668

detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual  
5 overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

10 Methods which may also be used to quantify the expression of TPPT include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of  
15 interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large  
20 numbers of genes simultaneously as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and  
25 monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

30 In another embodiment, antibodies specific for TPPT, or TPPT or fragments thereof may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci.  
35 USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-

WO 00/78953

PCT/US00/16668

2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding TPPT may be used  
 5 to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a  
 10 chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop  
 15 genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, e.g., Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic  
 20 map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding TPPT on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as  
 25 linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely  
 30 localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

35 In another embodiment of the invention, TPPT, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug

WO 00/78953

PCT/US00/16668

screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between TPPT and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with TPPT, or fragments thereof, and washed. Bound TPPT is then detected by methods well known in the art. Purified TPPT can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding TPPT specifically compete with a test compound for binding TPPT. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with TPPT.

In additional embodiments, the nucleotide sequences which encode TPPT may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, in particular U.S. Ser. No. 60/139,923, U.S. Ser. No. 60/148,177, U.S. Ser. No. 60/149,357, and U.S. Ser. No. 60/162,287, are hereby expressly incorporated by reference.

## EXAMPLES

### I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated





WO 00/78953

PCT/US00/16668

as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:44-86. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

#### IV. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum} \{ \text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}) \}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding TPPT occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous,

WO 00/78953

PCT/US00/16668

reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

#### V. Chromosomal Mapping of TPPT Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:44-49 and SEQ ID NO:54-86 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:44-49 and SEQ ID NO:54-86 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

The genetic map locations of SEQ ID NO:44, SEQ ID NO:48, SEQ ID NO:60, SEQ ID NO:65, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:73, SEQ ID NO:76, SEQ ID NO:80, and SEQ ID NO:83 are described in The Invention as ranges, or intervals, of human chromosomes. More than one map location is reported for SEQ ID NO:65, SEQ ID NO:73, SEQ ID NO:80, and SEQ ID NO:83, indicating that previously mapped sequences having similarity, but not complete identity, to SEQ ID NO:65, SEQ ID NO:73, SEQ ID NO:80, and SEQ ID NO:83 were assembled into their respective clusters. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Diseases associated with the public and Incyte sequences located within the indicated intervals are also reported in the Invention section where applicable. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

#### VI. Extension of TPPT Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:44-86 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this



WO 00/78953

PCT/US00/16668

fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target  
5 sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR  
10 was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing  $Mg^{2+}$ ,  $(NH_4)_2SO_4$ , and  $\beta$ -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C,  
15 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100  $\mu$ l PICOGREEN  
20 quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5  $\mu$ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5  $\mu$ l to 10  $\mu$ l aliquot of the reaction mixture was analyzed by  
25 electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For  
30 shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on  
35 antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was  
5 quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems).

10 In like manner, the polynucleotide sequences of SEQ ID NO:44-86 are used to obtain 5' regulatory sequences using the procedure above, along with oligonucleotides designed for such extension, and an appropriate genomic library.

#### VII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:44-86 are employed to screen cDNAs,  
15 genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase  
20 (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10<sup>7</sup> counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

25 The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and  
30 compared.

#### VIII. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra),  
mechanical microspotting technologies, and derivatives thereof. The substrate in each of the  
35 aforementioned technologies should be uniform and solid with a non-porous surface (Skena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers.

WO 00/78953

PCT/US00/16668

Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g.,  
 5 Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The  
 10 array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of  
 15 complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

#### Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and  
 20 poly(A)<sup>+</sup> RNA is purified using the oligo-(dT) cellulose method. Each poly(A)<sup>+</sup> RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)<sup>+</sup> RNA with  
 25 GEMBRIGHT kits (Incyte). Specific control poly(A)<sup>+</sup> RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37 °C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85 °C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc.  
 30 (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μl 5X SSC/0.2% SDS.

#### Microarray Preparation

35 Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification

WO 00/78953

PCT/US00/16668

uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 µg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

5 Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C  
10 oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 µl of the array element DNA, at an average concentration of 100 ng/µl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

15 Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60 °C followed by washes in 0.2% SDS and distilled water as before.

#### 20 Hybridization

Hybridization reactions contain 9 µl of sample mixture consisting of 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65 °C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm<sup>2</sup> coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly  
25 larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 µl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60 °C. The arrays are washed for 10 min at 45 °C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45 °C in a second wash buffer (0.1X SSC), and dried.

#### Detection

30 Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-  
35 scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

#### **IX. Complementary Polynucleotides**

Sequences complementary to the TPPT-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring TPPT. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of TPPT. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the TPPT-encoding transcript.

#### **X. Expression of TPPT**

Expression and purification of TPPT is achieved using bacterial or virus-based expression systems. For expression of TPPT in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express TPPT upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of TPPT in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding TPPT by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, TPPT is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from TPPT at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified TPPT obtained by these methods can be used directly in the assays shown in Examples XI and XV.

#### **XI. Demonstration of TPPT Activity**

TPPT transport activity is assayed by measuring uptake of labeled substrates into Xenopus laevis oocytes. Oocytes at stages V and VI are injected with TPPT mRNA (10 ng per oocyte) and incubated for 3 days at 18°C in OR2 medium (82.5mM NaCl, 2.5 mM KCl, 1mM CaCl<sub>2</sub>, 1mM MgCl<sub>2</sub>, 1mM Na<sub>2</sub>HPO<sub>4</sub>, 5 mM Hepes, 3.8 mM NaOH, 50µg/ml gentamycin, pH 7.8) to allow expression of TPPT. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM CaCl<sub>2</sub>, 1mM MgCl<sub>2</sub>, 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g., amino acids,

WO 00/78953

PCT/US00/16668

sugars, drugs, ions, and neurotransmitters) is initiated by adding labeled substrate (e.g. radiolabeled with  $^3\text{H}$ , fluorescently labeled with rhodamine, etc.) to the oocytes. After incubating for 30 minutes, uptake is terminated by washing the oocytes three times in  $\text{Na}^+$ -free medium, measuring the incorporated label, and comparing with controls. TPPT activity is proportional to the level of internalized labeled substrate.

## XII. Functional Assays

TPPT function is assessed by expressing the sequences encoding TPPT at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10  $\mu\text{g}$  of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2  $\mu\text{g}$  of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of TPPT on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding TPPT and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding TPPT and other genes of interest can be analyzed by northern analysis or microarray techniques.

### **XIII. Production of TPPT Specific Antibodies**

TPPT substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

5 Alternatively, the TPPT amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

10 Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (PE Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for anti-peptide and anti-TPPT activity by, for example, binding the peptide or TPPT to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

### **XIV. Purification of Naturally Occurring TPPT Using Specific Antibodies**

Naturally occurring or recombinant TPPT is substantially purified by immunoaffinity chromatography using antibodies specific for TPPT. An immunoaffinity column is constructed by covalently coupling anti-TPPT antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing TPPT are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of TPPT (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/TPPT binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and TPPT is collected.

### **XV. Identification of Molecules Which Interact with TPPT**

TPPT, or biologically active fragments thereof, are labeled with  $^{125}\text{I}$  Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled TPPT, washed, and any wells with labeled TPPT complex are assayed. Data obtained using different concentrations of TPPT are used to calculate values for the number, affinity, and association of TPPT with the candidate molecules.

35 Alternatively, molecules interacting with TPPT are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, *Nature* 340:245-246), or using commercially





Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	44	264114	HNT2AGT01	028972R6 (SPLNFET01), 028972T6 (SPLNFET01), 264114H1 (HNT2AGT01), 452387R6 (TYMNOT02), 736580R1 (TONSNOT01), 747955R6 (BRAITUT01), 936731R1 (CERVNOT01), 320628H1 (PENCNOT03), 3344943H1 (SPLNNOT09), 3742964H1 (THYMNOT08), 4028320H1 (BRAINOT23), 4726757H1 (GBLADIT01), 5473562H1 (MCLRUNT01)
2	45	1455669	COLNFET02	1455669H1 (COLNFET02), 2877376F6 (THYRNOT10), 3536452F6 (KIDNNOT25)
3	46	2084989	PANCNOT04	1281527H1 (COLNNOT16), 1412985H1 (BRAINOT12), 2084989H1 (PANCNOT04), 2084989R6 (PANCNOT04), 2084989T6 (PANCNOT04), 2470481F6 (THP1NOT03), 2539015F7 (BONRTUT01), 3109754F6 (BRSTTUT15), 3694831H1 (PANCNOT19), 3700647H1 (SININOT05)
4	47	2501034	ADRETUT05	111466F1 (PITUNOT01), 111466R1 (PITUNOT01), 414042R6 (BRSTNOT01), 687891H1 (UTRSNOT02), 2501034H1 (ADRETUT05)
5	48	2745212	LUNGCTUT11	000802H1 (U937NOT01), 008963H1 (HMC1NOT01), 009314H1 (HMC1NOT01), 135428F1 (BMARNOT02), 723168X19 (SYNCOAT01), 1000842R1 (BRSTNOT03), 1370189H1 (BSTMNON02), 1374329H1 (BSTMNON02), 2745212H1 (LUNGCTUT11), 4920466H1 (TESTNOT11), SAIA02182F1
6	49	4833111	BRAVTXT03	864776T1 (BRAITUT03), 1911267F6 (CONNTUT01), 4833111H1 (BRAVTXT03), SARA02608F1, SARA02002F1
7	50	876677	LUNGAST01	876677H1 (LUNGAST01), 876677R6 (LUNGAST01), SCDA08642V1
8	51	2326143	OVARNOT02	867305R1 (BRAITUT03), 963058R2 (BRSTTUT03), 1715155F6 (UCMCNOT02), 1727927T6 (PROSNOT14), 2326143H1 (OVARNOT02), 2326143R6 (OVARNOT02), 3360563H1 (PROSTUT16)
9	52	2786302	BRSTNOT13	2786302H1 (BRSTNOT13), 2958321X303D1 (ADRENOT09), 2958321X305D1 (ADRENOT09), 2958321X308D1 (ADRENOT09)
10	53	3735780	SMCCNOS01	551126H1 (BEPINOT01), 2808373H1 (BLADTUT08), 3735780F6 (SMCCNOS01), 3735780H1 (SMCCNOS01), 3735780T6 (SMCCNOS01), 4760604T6 (BRAMNOT01)
11	54	039026	HUVENOB01	039026H1 (HUVENOB01), 159164F1 (ADENINB01), 159164R1 (ADENINB01)
12	55	260607	HNT2RAT01	063159R6 (PLACNOB01), 260607R6 (HNT2RAT01), 1272850T1 (TESTTUT02), 1273069H1 (TESTTUT02), 2867453F6 (KIDNNOT20), 3082466H1 (BRAIUNT01), 4796739H1 (LIVRTUT09), 4799318F6 (MYEPUNT01), g1424405

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
13	56	1429651	SINTBST01	1429651F1 (SINTBST01), 1429651H1 (SINTBST01), 1501096F6 (SINTBST01), 1989621T6 (CORPNOT02), SXLA01343V1, SXLA01183V1, SXLA01559V1, SXLA00812V1
14	57	2069971	ISLTNOT01	2069606F6 (ISLTNOT01), 2069971H1 (ISLTNOT01), 2374634F6 (ISLTNOT01), 2383754F6 (ISLTNOT01), 4171186T6 (SINTNOT21), SXLA01128V1, SXLA01348V1, SXLA01219V1, SXLA00260V1, SXLA00074V1
15	58	2329339	COLNNOT11	658662H1 (BRAINOT03), 1544110R1 (PROSTUT04), 1657742F6 (URETTUT01), 1750523F6 (STOMTUT02), 2329339H1 (COLNNOT11), 2329339R6 (COLNNOT11), 3858671H1 (LNODNOT03), g1494061, g1891451
16	59	2540219	BONRTUT01	2540219H1 (BONRTUT01), 2540219T6 (BONRTUT01), 2554869F6 (THYMNOT03), g869197
17	60	2722462	LUNGTUT10	883601R1 (PANCNOT05), 1525902F6 (UCMCL5T01), 1525902X18C1 (UCMCL5T01), 1525902X311D1 (UCMCL5T01), 1527325T6 (UCMCL5T01), 1554770X311D1 (BLADTUT04), 2417265H1 (HNT3AZT01), 2444786F6 (THP1NOT03), 2722462H1 (LUNGTUT10), 4293114H1 (BRABDIR01), 5070268T6 (PANCNOT23), SANA01850F1, SAJA01078R1, SANA02081F1, SAJA01813F1
18	61	2739264	OVARNOT09	000573H1 (U937NOT01), 494409F1 (HNT2NOT01), 494409R1 (HNT2NOT01), 2506506F6 (CONUTUT01), 2681059H1 (SINIUCT01), 2744648F6 (BRSTTUT14), 2805590F6 (BLADTUT08), 3770643H1 (BRSTNOT25), 4204278H1 (BRAITUT29), SAEA02093F1
19	62	2758310	THPLAZS08	487309R7 (HNT2AGT01), 1361439F1 (LUNGNOT12), 2758310H1 (THP1AZS08), SCFA05584V1, SCFA05940V1, SCFA05166V1, SCFA05135V1
20	63	2762348	BRSTNOT12	632097R6 (KIDNNOT05), 632097T6 (KIDNNOT05), 2762348H1 (BRSTNOT12), SCCA02837V1, SCCA05356V1, SCCA01377V1, SCCA05963V1, SCCA05364V1, SCCA02307V1, SCCA04327V1, SCCA02009V1
21	64	3715961	PENCNOT09	961523H1 (BRSTTUT03), 1863723F6 (PROSNOT19), 2265329H1 (UTRSNOT02), 2360619R6 (LUNGFET05), 2360619T6 (LUNGFET05), 2821718H1 (ADRETUT06), 3715961H1 (PENCNOT09), 5016160H1 (BRAXNOT03), 5499583H1 (BRABDIR01)
22	65	5108194	PROSTUS19	1322651X35 (BLADNOT04), 1322651X36 (BLADNOT04), 3494841H1 (ADRETUT07), 4958978F6 (TLYMNOT05), 5108194H1 (PROSTUS19), g1379009, g1527417
23	66	5503122	BRABDIR01	5503122F6 (BRABDIR01), 5503122H1 (BRABDIR01), 5503122R6 (BRABDIR01)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
24	67	5517972	LIVRDIR01	805957R1 (BSTNOT01), 953622R1 (SCORN01), 1501080F1 (SINTBST01), 1547381R6 (PROSTUT04), 2081843T6 (UTRSNOT08), 2811524F6 (OVARNOT10), 3212921H1 (BLADNOT08), 3250443H1 (SEMVN03), 3269479H1 (BRAINT02), 3699955F6 (SININOT05), 3700568H1 (SININOT05), 4944050H1 (BRAIFEN05), 5517972H1 (LIVRDIR01)
25	68	5593114	COLCDIT03	2859465F6 (SININOT03), 2859465T6 (SININOT03), 3555656F6 (LUNGNOT31), 3555656T6 (LUNGNOT31), 4345952H1 (TLYMTXT01), 5593114H1 (COLCDIT03), 5874544H1 (COLTDIT04)
26	69	044775	TBLYN0T01	044775H1 (TBLYN0T01), 044775X3 (TBLYN0T01), 455640R1 (KERANOT01), 950702R1 (PANCNOT05), 2418550H1 (HNT3AZT01), 2798917H1 (NPOLNOT01), 2844696H1 (DRGLNOT01), g1718929
27	70	116588	KIDNNOT01	699714R6 (SYNORAT03), 831423R1 (PROSTUT04), 978875R1 (BRSTNOT02), 1350569F1 (LATRTUT02), 1447681R1 (PLACNOT02), 3177382F6 (UTRSTUT04), 3688796H1 (HEAANOT01), 3929008H1 (KIDNNOT19), g2106455, g2163092
28	71	875369	LUNGAST01	571573F1 (OVARNON01), 571573R1 (OVARNON01), 875369H1 (LUNGAST01), 875369R1 (LUNGAST01), 3569021H1 (HEAPNOT01)
29	72	1325518	LPARNOT02	1325518H1 (LPARNOT02), 1325518T6 (LPARNOT02), 1825553F6 (LSUBNOT03), SBAA02035F1
30	73	2060987	OVARNOT03	1378947T1 (LUNGNOT10), 1453290F1 (PENITUT01), 1459818R1 (COLNFET02), 1967477H1 (BRSTNOT04), 2060987H1 (OVARNOT03), 2455371F6 (ENDANOT01), 2499967F7 (ADRETUT05), 3093056T6 (BRSTNOT19), 3213366H1 (BLADNOT08), 4934158H1 (BRSTTUT20), SBYA01942U1
31	74	2172064	ENDCNOT03	2172064CT1 (ENDCNOT03), 2172064H1 (ENDCNOT03), SBLA01269F1
32	75	2219267	LUNGNOT18	2219267F6 (LUNGNOT18), 2219267H1 (LUNGNOT18), 3117478T6 (LUNGNOT13), 3126288T6 (LUNGNOT12), 3558495H1 (LUNGNOT31)
33	76	2308629	NGANNOT01	469862F1 (MMLRIDT01), 469862R1 (MMLRIDT01), 1594203X11C1 (BRAINT04), 2191933H1 (THYRTUT03)
34	77	2660038	LUNGTUT09	1326594F1 (LPARNOT02), 2256143H1 (OVRTUT01), 2278689R6 (PROSNON01), 2528425H1 (GBLANOT02), 2660038H1 (LUNGTUT09), 2660038T6 (LUNGTUT09), 3449964H1 (UTRSNON03), 5099879H1 (PROSTUS20), g1886680, g783969

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
35	78	2670745	ESOGTUT02	259200X12 (HNT2RAT01), 1266477F1 (BRAINOT09), 2383364F6 (ISLTNOT01), 2670745H1 (ESOGTUT02), 3181526H1 (TLYJNOT01)
36	79	2676443	KIDNNOT19	607375R6 (BRSTTUT01), 1728626X15C1 (PROSNOT14), 1751773F6 (LIVRTUT01), 1751994T6 (LIVRTUT01), 1796032X14C1 (PROSTUT05), 2010172H1 (TESTNOT03), 2676443H1 (KIDNNOT19)
37	80	3295764	TLYJINT01	063264H1 (PLACNOB01), 434468T6 (THYRNOT01), 487721H1 (HNT2AGT01), 907796R2 (COLNNOT09), 1212556R7 (BRSTTUT01), 1251889H1 (LUNGFET03), 1653370F6 (PROSTUT08), 1653370X309D1 (PROSTUT08), 2192762F6 (THYRTUT03), 2226786F6 (SEMVNOT01), 3295764H1 (TLYJINT01), 3384471H1 (ESOGNOT04), SASA01137F1
38	81	3438320	PENCNOT06	3438320H1 (PENCNOT06), 3501438H1 (PROSTUT13), 3745542H1 (THYMNOT08), 3751060H1 (UTRSNOT18), 4979750F6 (HELATXT04), SADA00043F1, SADA00087F1
39	82	3986488	UTRSTUT05	1634141F6 (COLNNOT19), 1692115X12C1 (PROSTUT10), 1731310F6 (BRSTTUT08), 2046232H1 (THP1T7T01), 3557951H1 (LUNGNOT31), 4726788H1 (GBLADIT01)
40	83	4378816	LUNGNOT37	1318962H1 (BLADNOT04), 1520864F1 (BLADTUT04), 1684381F6 (PROSNOT15), 2055747R6 (BEPINOT01), 4378816H1 (LUNGNOT37)
41	84	4797137	LIVRTUT09	4797137F6 (LIVRTUT09), 4797137H1 (LIVRTUT09), 4797137T6 (LIVRTUT09)
42	85	5470806	MCLRUNT01	5470806H1 (MCLRUNT01), 5470806T6 (MCLRUNT01)
43	86	5473242	MCLRUNT01	5473242F6 (MCLRUNT01), 5473242T6 (MCLRUNT01)

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
1	623	S521 S2 T3 S16 S99 S138 S144 T193 T264 T404 S448 S589 S151 T229 T337 T457 S562 S568	N97 N333	BTB domain: C44-F56 POZ domain: N10-Q211 Kelch repeat signature: E379-G392, T398-V412, L438-M452, T498-A512 Ring canal protein repeat: E122-P254	Ring canal protein [Drosophila melanogaster] g577276	MOTIFS BLIMPS-PFAM BLIMPS-PRINTS BLAST-GenBank BLAST-PRODOM BLAST-DOMO
2	99	T17	N15	Signal peptide: M1-G36 Transmembrane region: S25-W45 MRP(2) MRP(1) repeat: C30-V74	Multi-drug resistance- associated protein- (MRP)-like protein- 1 (MLP-1) [Rattus norvegicus] g3242458	MOTIFS BLAST-GenBank BLAST-PRODOM SPScan HMMER
3	374	T334 T33 S137 T146 S291 S311 T346	N103 N127 N135 N138	Signal peptide: M1-N52	Tricarboxylate carrier [Rattus sp.] g545998	MOTIFS BLAST-GenBank SPScan
4	271	S234 T126 T169 Y141		Signal peptide: M1-C30 Transmembrane region: L233-F252	Weak similarity with honeybee ATP synthase A chain [Caenorhabditis elegans] g3878801	MOTIFS BLAST-GenBank SPScan HMMER
5	323	S99 S125 S192 T277 S307 S309 T110 Y212		Leucine zipper: L284-L305	Cu <sup>2+</sup> -transporting ATPase homolog [Arabidopsis thaliana] g2464854	BLAST-GenBank MOTIFS
6	274	S96 T198 S215 T29 S121 S164 S170		Mitochondrial energy transfer proteins: G5-L266 Signal peptide: M1-G17	Pet8p [Saccharomyces cerevisiae] g495307	BLAST-GenBank HMMER-PFAM MOTIFS ProfilesScan BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO SPScan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
7	291	S6 T113 T173 T147 S230 T258	N226 N261	Signal peptide: M1-T42 Transmembrane domain: W29-I54 Band 7 protein family: C50-V62, K90-E224 Membrane stomatin: E14-N283	Stomatin [Homo sapiens] g1161562	MOTIFS BLAST-GenBank SPScan HMMER BLIMPS-BLOCKS BLAST-DOMO BLAST-PRODOM
8	381	S2 S35 T57 S92 T104 S191 S302 S334 S335 S336 T43 T250 T255 T304 S311 S370 Y65	N218 N253 N259		K <sup>+</sup> channel modulatory factor DEBT-91 [Mus musculus] g4838557	MOTIFS BLAST-GenBank
9	190	T160 S17 T71 S77 T78 S111 S134 S142	N87	ABC transporter family: R79-K177 ATP/GTP-binding site motif A (P-loop): G102-S109	ABC2 transporter [Mus musculus] g495259	MOTIFS BLAST-GenBank BLAST-DOMO
10	297	S17 S114 T136 S16	N287	Mitochondrial carrier protein signature: E117-I297 Graves Disease carrier protein: P137-T157, L259-S279	Similar to human ADP/ATP carrier protein [C. elegans] g3879938	MOTIFS BLAST-GenBank HMMER-PFAM BLIMPS-PRINTS
11	89	T37 T47 T60 S64			Mitochondrial import protein Tim9p [Saccharomyces cerevisiae] g3747026	BLAST-GenBank MOTIFS
12	115	T108 T84		Signal peptide: M1-G24 Transmembrane domain: G35-F57 Sodium neurotransmitter symporter signature: R7-S61		MOTIFS SPScan HMMER ProfileScan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
13	675	T54 T50 S99 T127 S413 T558 S645 T654 T47 S242 T602 T611 Y501	N243 N247 N301 N601	Transmembrane domains: I29-V48, L103-I121, L177-G196, I210-M229, L417-W435, F481-Y501, Y521-W541 Sodium symporter family domain: Y58-G487 Sodium:solute symporter signature: Y35-G89, M111-R140, L173-G227, P460-G469	Sodium-glucose cotransporter [Oryctolagus cuniculus] g473969	BLAST-GenBank MOTIFS HMMER HMMER-PFAM BLIMPS-BLOCKS ProfileScan BLAST-PRODOM BLAST-DOMO
14	320	T84 S304 T11 S75 S80 S164 Y20	N162 N234	Transmembrane domains: I92-L112, I201-K219 Zinc transporter signature: A28-V142, D199-E303 Cation transporter domain: S48-L74	Zinc transporter ZnT-2 [Rattus norvegicus] g1256378	BLAST-GenBank MOTIFS HMMER BLIMPS-PRODOM BLAST-PRODOM BLAST-DOMO
15	462	S111 S145 S183 S233 T26 T185 S202 T243	N24 N279	Kelch repeat motifs: C299-N349; F350-R399 Y400-G446 BTB domain: F50-L117 POZ domain: Y27-E215	Ring canal protein [Drosophila melanogaster] g577276	BLAST-GenBank MOTIFS HMMER-PFAM BLIMPS-PRINTS BLAST-DOMO
16	98	T22 Y37		Signal peptide: M1-S17 Mitochondrial carrier proteins domain: C4-I89 Mitochondrial carrier proteins signature sequence: V6-G19, G19-A33, G63-E83	Carrier protein (cl) [Caenorhabditis elegans] g472902	BLAST-GenBank MOTIFS SPScan HMMER-PFAM ProfileScan BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-DOMO



Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
17	748	S55 S196 T254 S307 S327 T491 T534 T550 T571 S635 S648 S677 T696 S283 S291 T314 S629 S701 Y556	N531 N543 N548 N627	Signal peptide: M1-A61 Transmembrane domains: L39-L56, I167-F186, C229-F252, G438-L455, M492-F509, L598-I618 Ion transport proteins signature: F85-V251, L369-I618	Voltage-gated calcium channel [Rattus norvegicus] g4586963	BLAST-GenBank MOTIFS SPScan HMMER HMMER-PFAM BLIMPS-PRINTS
18	507	T200 S183 T232 T284 T349 T150 T252 S253 S319 S383 Y454	N220 N250 N364 N496	Signal peptide: M1-G26	Nucleoporin p54 [Rattus norvegicus] g1537070	BLAST-GenBank MOTIFS SPScan
19	592	S460 S104 T178 S320 S321 T498 T531 Y365		ABC1 precursor signature: N153-Q162, F210-A229, G234-I254, V312-G332, T366-V378	ABC transporter [Methanobacterium thermo.] g2622773	BLAST-GenBank MOTIFS BLIMPS-PRODOM BLAST-PRODOM
20	841	T98 S120 S203 T214 T276 S388 T438 T700 T838 T167 T179 S280 T370 S435 S531 S539 S666 S693 S830	N368 N490 N624	Transmembrane domains: Y451-D469, M544-F562, F577-F597, G775-M797 Vacuolar ion transport subunit signature: M10-F831	Vacuolar H <sup>+</sup> /ATPase subunit [Rattus norvegicus] g206430	BLAST-GenBank MOTIFS HMMER BLIMPS-PRODOM BLAST-PRODOM BLAST-DOMO
21	253	S50 T139 T152 T177 S202 T143 Y55		Mitochondrial carrier proteins domain: Y31-S248 Mitochondrial energy transfer proteins signature sequence: I62-O86, I110-G122	Mitochondrial uncoupling protein UCP-4 [Homo sapiens] g4324701	BLAST-GenBank MOTIFS HMMER-PFAM BLIMPS-BLOCKS ProfileScan BLAST-PRODOM BLAST-DOMO

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
22	229	S69 S26 S109 T162 S178 S25 S64 S65 T210 S219		Signal peptide: M1-A47 Mitochondrial carrier proteins domain: Q32-G220 Mitochondrial carrier proteins signature sequence: S36-T49, T49-V63, G92-E112, T144-T162, Y187-F205	Grave's disease carrier protein [Bos taurus] g387	BLAST-GenBank MOTIFS SPScan HMMER-PFAM BLIMPS-BLOCKS ProfileScan BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO
23	170	S26 S31 S149 S164 T22 T157	N66 N145	Dihydroxyipyridine-sensitive L-type calcium channel signature: Y2-A47, I49-V77, A83-N100, R106-E131 SH3 domain: V59-R122	Voltage-dependent calcium channel beta-4 subunit [Homo sapiens] g2058727	BLAST-GenBank MOTIFS HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLIMPS-PFAM BLAST-PRODOM BLAST-DOMO
24	655	T194 S195 S232 T362 S655 S4 S88 T135 T153 S187 T214 S322 T345 S353 S443 T609 S261 S381 S384	N338 N418 N557 N596	Transmembrane domains: I396-K417, Y494-S522, T538-V556 ABC transporters domain: P73-G262 ABC transporter family signature sequence: I78-L89, V186-D217	Breast cancer resistance protein (multidrug transporter) [Homo sapiens] g4038352	BLAST-GenBank MOTIFS HMMER HMMER-PFAM BLIMPS-BLOCKS ProfileScan BLAST-PRODOM BLAST-DOMO
25	184	T51 S29 T100 S138 S151 Y78	N27		Cation transport protein [E. coli] g495778	BLAST-GenBank MOTIFS
26	154	S54 S42 S62 T78 Y104		Mitochondrial energy transfer proteins signatures: P89-L97, M1-E41, M73-L152 Mitochondrial carrier protein domain: G2-L152	Similar to carrier protein C2 [C. elegans] g3879669	MOTIFS HMMER-PFAM BLAST-PRODOM BLAST-DOMO BLAST-GenBank

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
27	438	S170 T5 T51 T265 T300 S425	N50 N423	Transmembrane domains: C91-L111, L237-I257, I305-M332, M332-L354, L391-V408, I186-A204	Multidrug efflux transporter [Bacillus subtilis] g2635104	MOTIFS HMMER BLAST-GenBank
28	237	S10 S47 T72 S28 S96 S148 T173 T222 S6 S21 T32 T61 T192	N35	Nucleic acid-binding protein E5.1 domain: S6-K128	ARL-6 interacting protein-4 [Mus musculus] g4927204	MOTIFS BLAST-DOMO BLAST-GenBank
29	219	T66 S194 T200		Signal peptide: M1-R19 or M1-K15 Caseins alpha/beta signature: M1-N39	Surface antigen [Trypanosoma cruzi] g161956	MOTIFS HMMER SPScan ProfileScan BLAST-GenBank
30	707	S31 T6 T55 T263 T328 T546 T580 T594 S662 S673 T32 S50 S231 T244 T306 T385 S439 S476 S533 S553 S624	N343 N570 N638 N703	Potassium channel signature: A62-T81 Potassium channel integral membrane protein domain: S13-D117	NY-REN-45 antigen (similar to potassium channel protein) [Homo sapiens] g5360115	MOTIFS BLIMPS-PRINTS BLAST-DOMO BLAST-GenBank
31	279	T18 T245 T206	N181	Signal cleavage: M1-G45 Connexin domains: M1-V99, V20-Y44 Connexin signatures: L33-V86, L152-F205, F51-P73, S76-L96, L133-Y159, C169-T189, I190-L218 Gap junction protein connexin transmembrane regions: F5-Y97, L133- K223, M1-S130	Gap junction protein (similar to connexin) [Homo sapiens] g3006230	MOTIFS SPScan HMMER BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PFAM ProfileScan BLAST-PRODOM BLAST-DOMO BLAST-GenBank

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
32	154	S114		Signal peptide: M1-A35 or M1-A14 Transmembrane domain: F83-L102	mBOCT (potent organic cation transporter) [Mus musculus] g4589468	MOTIFS HMME SPScan BLAST-GenBank
33	289	T83 T205 S269 T279	N60	Mitochondrial energy transfer proteins signatures: M1-G147, P17, P115, N185-K280, A101-Q181, Y184-I278 Mitochondrial carrier protein domains: M1-E176, N185-K280 Mitochondrial transmembrane transport protein regions: P17-R182, P180-I278	Mitochondrial solute carrier [Onchocerca volvulus] g1518458	MOTIFS HMME-PFAM BLAST-DOMO BLAST-PRODOM ProfileScan BLAST-GenBank
34	300	S189 S195 S204 T257		Mitochondrial energy transfer proteins signatures: P19-M27, D2-I53, L209-L295 Mitochondrial carrier protein domain: D2-Y295 Transport protein domain: P122-Y295	YKL522=mitochondria 1 ADP/ATP carrier protein homolog [Saccharomyces cerevisiae] g254449	MOTIFS HMME-PFAM ProfileScan BLAST-PRODOM BLAST-DOMO BLAST-GenBank
35	382	S34 S207 T221 S312 T40 S53 T112 T117 T277 S337	N96 N372	Kelch motifs: H191-G249, E250-D301	Similarity to Human host cell factor C1 [Homo sapiens] g3875291	MOTIFS HMME BLAST-PFAM BLAST-GenBank

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
36	287	T36 T118 S180 S230 T84 S168 T244		Mitochondrial energy transfer proteins signatures: P26-L34, P219-L227, L97-G193, W10-V89, D197-F281, P96-Y194 Mitochondrial carrier protein domain: A5-F281 Mitochondrial brown fat region: Y82-Q94, V151-S168, Y194-C212	Mitochondrial dicarboxylate carrier [Rattus norvegicus] g3646426	MOTIFS HMMER-PFAM BLIMPS-PRINTS BLAST-DOMO BLAST-GenBank
37	497	T65 T135 S147 T360 S8 T22 S45 S291	N63 N314 N414	Transmembrane domains: M114-T137, M364-M380, Y390-A413, A421-D444, F456-V478 Folate transporter domains: W30-H218, I253-K484	Reduced folate carrier [Homo sapiens] g1041934	MOTIFS HMMER BLAST-PRODOM BLAST-DOMO SPScan BLAST-GenBank
38	228	T21 S124 T145 S158 T190 T95 S132 S137 T177		Heme-binding domain in cytochrome b5: Y19-G98 Cytochrome b5 family domain: H28-P75	cytochrome b5 containing fusion protein [Helianthus annuus] g1040729 P=1.2e-07	MOTIFS HMMER-PFAM BLAST-GenBank ProfileScan
39	273	T63 S158 T48	N214	Transmembrane domains: L85-N105, F180-Y200 Intermembrane space domain: L30-L251	Sqv-7-like protein (similar to nucleotide-sugar transporters) [Homo sapiens] g4008517	MOTIFS HMMER BLAST-DOMO BLAST-GenBank

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
40	206	S187 S201	N158	Signal peptide: M1-G29 or M1-A27 Emopamil binding protein: G37-S187, L15-K203 Transmembrane domain: Y164-L183	C-8,7 sterol isomerase, aSI1 [Arabidopsis thaliana] g2772934	MOTIFS HMMER ProfileScan BLAST-DOMO BLAST-GenBank
41	235	S192 S200 S56 T95 T146 S199 T207 S229 T53 T61 T69 T119 T148 Y70	N123	Transmembrane domain: F15-I34, M155-V174 Channel myelin protein: L18-M181 Sodium channel beta-2 subunit precursor: F15-E210 Immunoglobulin domain: I34-V136	Myelin protein zero (MPZ) [Homo sapiens] g2160399	MOTIFS HMMER BLIMPS-PRINTS BLAST-PRODUM BLAST-DOMO BLAST-GenBank
42	147	T79 T116 S3 S66 Y89 Y98	N118	Signal peptide: M1-G23 or M1-A20 Transthyretin signature: S28-S132 Transthyretin domain: G21-Q146	Transthyretin precursor [Sus scrofa] g1009702	MOTIFS HMMER ProfileScan BLAST-PRODUM BLAST-DOMO BLAST-GenBank BLIMPS-BLOCKS BLIMPS-PRINTS
43	147	T5 S88 T39		Globin domain: V2-H147 Heme oxygen transport protein domain: L32-H147	III beta-3 globin [Rattus norvegicus] g395943	MOTIFS HMMER-PFAM BLAST-PRODUM BLAST-DOMO BLIMPS-BLOCKS BLIMPS-PRINTS

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
44	1567-1611 2107-2151	Gastrointestinal (0.203) Hematopoietic/Immune (0.188) Nervous (0.156)	Cell Proliferation and Cancer (0.547) Inflammation (0.422)	PBLUESCRIPT
45	1-92 351-434	Endocrine (0.333) Developmental (0.167) Gastrointestinal (0.167) Musculoskeletal (0.167) Reproductive (0.167)	Cell Proliferation and Cancer (0.833) Inflammation (0.167)	pINCY
46	920-964 1352-1396	Reproductive (0.304) Gastrointestinal (0.174) Cardiovascular (0.130) Hematopoietic/Immune (0.130) Nervous (0.130)	Cell Proliferation and Cancer (0.478) Inflammation (0.391)	PSPORT1
47	1-80 768-848	Nervous (0.273) Reproductive (0.273) Gastrointestinal (0.127) Hematopoietic/Immune (0.127)	Cell Proliferation and Cancer (0.564) Inflammation (0.400)	pINCY
48	111-194 687-758	Reproductive (0.221) Nervous (0.185) Gastrointestinal (0.124)	Cell Proliferation and Cancer (0.552) Inflammation (0.343)	pINCY
49	1-97	Nervous (0.234) Hematopoietic/Immune (0.191) Gastrointestinal (0.149)	Cell Proliferation and Cancer (0.617) Inflammation (0.340)	pINCY
50	218-262	Cardiovascular (1.000)	Cancer (0.333) Inflammation/Trauma (0.333) Cell Proliferation (0.333)	PSPORT1
51	811-855	Hematopoietic/Immune (0.180) Gastrointestinal (0.146) Reproductive (0.281)	Cancer (0.393) Inflammation/Trauma (0.515) Cell Proliferation (0.146)	PSPORT1
52	595-639	Gastrointestinal (0.286) Reproductive (0.714)	Cancer (0.429) Inflammation/Trauma (0.429)	pINCY
53	96-140	Cardiovascular (0.167) Hematopoietic/Immune (0.167) Nervous (0.250) Reproductive (0.167)	Cancer (0.250) Inflammation/Trauma (0.167) Cell Proliferation (0.167)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
54	507-551	Reproductive (0.323) Gastrointestinal (0.154) Nervous (0.123)	Cancer (0.446) Inflammation/Trauma (0.308) Cell Proliferation (0.185)	PBLUESCRIPT
55	455-499	Urologic (0.333) Nervous (0.222) Reproductive (0.222)	Cancer (0.667) Cell Proliferation (0.333)	PBLUESCRIPT
56	1835-1879	Nervous (0.625) Gastrointestinal (0.375)	Inflammation/Trauma (0.375) Cancer (0.250) Neurological (0.250)	pINCY
57	811-855	Gastrointestinal (1.000)	Inflammation/Trauma (0.667)	pINCY
58	390-434	Reproductive (0.320) Nervous (0.240) Urologic (0.120)	Cancer (0.520) Inflammation/Trauma (0.240) Cell Proliferation (0.160)	PSPORT1
59	413-457	Gastrointestinal (0.333) Musculoskeletal (0.333) Nervous (0.333)	Cancer (0.333) Neurological (0.333)	pINCY
60	2021-2084	Nervous (0.197) Gastrointestinal (0.184) Reproductive (0.184)	Cancer (0.461) Inflammation/Trauma (0.316) Cell Proliferation (0.118)	pINCY
61	65-109	Nervous (0.226) Reproductive (0.208) Cardiovascular (0.113) Gastrointestinal (0.113)	Cancer (0.528) Inflammation/Trauma (0.301) Cell Proliferation (0.208)	pINCY
62	379-423 1867-1911	Reproductive (0.282) Gastrointestinal (0.205) Nervous (0.154)	Cancer (0.538) Inflammation/Trauma (0.282) Cell Proliferation (0.103)	PSPORT1
63	362-406 1193-1237	Urologic (0.500) Reproductive (0.333) Cardiovascular (0.167)	Cancer (0.667) Inflammation/Trauma (0.333)	pINCY
64	394-438	Nervous (0.294) Reproductive (0.265) Cardiovascular (0.118)	Cancer (0.382) Inflammation/Trauma (0.235) Cell Proliferation (0.118)	pINCY



Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
65	768-812	Reproductive (0.300) Endocrine (0.200) Gastrointestinal (0.200) Hematopoietic/Immune (0.200)	Inflammation/Trauma (0.500) Cancer (0.400)	pINCY
66	77-121	Nervous (1.000)	Neurological (1.000)	pINCY
67	1999-2043	Reproductive (0.324) Nervous (0.265) Gastrointestinal (0.235)	Cancer (0.500) Inflammation/Trauma (0.294) Cell Proliferation (0.118)	pINCY
68	561-605	Hematopoietic/Immune (0.455) Gastrointestinal (0.182) Nervous (0.182)	Inflammation/Trauma (0.546) Cell Proliferation (0.182)	pINCY
69	679-729	Nervous (0.292) Gastrointestinal (0.208) Hematopoietic/Immune (0.125)	Cancer (0.250) Cell Proliferation (0.375) Inflammation/Trauma (0.416)	PBLUESCRIPT
70	95-366 1078-1185	Reproductive (0.206) Hematopoietic/Immune (0.186) Cardiovascular (0.127)	Cancer (0.373) Inflammation/Trauma (0.382) Cell Proliferation (0.176)	PBLUESCRIPT
71	33-152	Reproductive (0.275) Nervous (0.163) Gastrointestinal (0.137)	Cancer (0.438) Inflammation/Trauma (0.314) Cell Proliferation (0.176)	PSPORT1
72	81-779	Gastrointestinal (1.000)	Cancer (1.000)	pINCY
73	719-817 1202-1414	Reproductive (0.311) Hematopoietic/Immune (0.203) Gastrointestinal (0.122)	Cancer (0.459) Inflammation/Trauma (0.379) Cell Proliferation (0.203)	PSPORT1
74	1-848	Nervous (0.750) Dermatologic (0.250)	Cancer (0.250) Cell Proliferation (0.250) Inflammation/Trauma (0.500)	pINCY
75	1-478	Cardiovascular (0.714) Developmental (0.143) Hematopoietic/Immune (0.143)	Cancer (0.571) Cell Proliferation (0.286) Inflammation (0.143)	pINCY
76	1-134	Reproductive (0.253) Nervous (0.241) Gastrointestinal (0.127) Hematopoietic (0.127)	Cancer (0.494) Inflammation (0.215) Cell Proliferation (0.127)	PSPORT1

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
77	510-719 960-1100	Reproductive (0.467) Cardiovascular (0.133) Gastrointestinal (0.133)	Cancer (0.467) Inflammation/Trauma (0.467)	pINCY
78	180-293	Reproductive (0.230) Nervous (0.225) Gastrointestinal (0.124)	Cancer (0.478) Inflammation/Trauma (0.292) Cell Proliferation (0.191)	pINCY
79	192-653 795-935	Reproductive (0.417) Gastrointestinal (0.292) Urologic (0.125)	Cancer (0.750) Cell Proliferation (0.125) Inflammation/Trauma (0.167)	pINCY
80	139-1044	Reproductive (0.245) Nervous (0.143) Developmental (0.122)	Cancer (0.490) Inflammation/Trauma (0.286) Cell Proliferation (0.224)	pINCY
81	233-916	Reproductive (0.667) Cardiovascular (0.167) Nervous (0.167)	Cancer (0.500) Cell Proliferation (0.333) Inflammation (0.167)	pINCY
82	1-153 760-816	Gastrointestinal (0.282) Hematopoietic/Immune (0.205) Reproductive (0.205)	Inflammation/Trauma (0.461) Cancer (0.308) Cell Proliferation (0.205)	pINCY
83	57-299	Nervous (0.179) Reproductive (0.179) Gastrointestinal (0.128)	Cancer (0.564) Cell Proliferation (0.256) Inflammation/Trauma (0.180)	pINCY
84	1-707	Gastrointestinal (0.500) Hematopoietic/Immune (0.500)	Cancer (0.500) Inflammation (0.500)	pINCY
85	451-594	Hematopoietic/Immune (1.000)	Cell Proliferation (1.000)	pINCY
86	8-124 161-187 407-472	Developmental (1.000)	Cell Proliferation (1.000)	pINCY

Table 4

SEQ ID NO:	Library	Library Comment
44	HNT2AGT01	Library was constructed at Stratagene (STR937233), using RNA isolated from the hNT2 cell line derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor. Cells were treated with retinoic acid for 5 weeks and with mitotic inhibitors for two weeks and allowed to mature for an additional 4 weeks in conditioned medium.
45	COLNFET02	Library was constructed using RNA isolated from the colon tissue of a Caucasian female fetus, who died at 20 weeks' gestation.
46	PANCNOT04	Library was constructed using RNA isolated from the pancreatic tissue of a 5-year-old Caucasian male, who died in a motor vehicle accident. Serology was positive for cytomegalovirus (CMV).
47	ADRETUT05	Library was constructed RNA isolated from adrenal tumor tissue removed from a 52-year-old Caucasian female during a unilateral adrenalectomy. Pathology indicated a pheochromocytoma.
48	LUNGTUT11	Library was constructed using RNA isolated from lung tumor tissue removed from the right lower lobe a 57-year-old Caucasian male during a segmental lung resection. Pathology indicated an infiltrating grade 4 squamous cell carcinoma. Multiple intrapulmonary peribronchial lymph nodes showed metastatic squamous cell carcinoma. Patient history included a benign brain neoplasm and tobacco abuse. Family history included spinal cord cancer, type II diabetes, cerebrovascular disease, and malignant prostate neoplasm.
49	BRAVXT03	Library was constructed using RNA isolated from treated astrocytes removed from the brain of a female fetus who died after 22 weeks' gestation. The cells were treated with tumor necrosis factor (TNF) alpha and interleukin 1 (IL-1), 10ng/ml each for 24 hours.
50	LUNGAST01	Library was constructed using RNA isolated from the lung tissue of a 17-year-old Caucasian male, who died from head trauma. Patient history included asthma.
51	OVARNOT02	Library was constructed using RNA isolated from ovarian tissue removed from a 59-year-old Caucasian female who died of a myocardial infarction. Patient history included cardiomyopathy, coronary artery disease, previous myocardial infarctions, hypercholesterolemia, hypotension, and arthritis.
52	BRSTNOT13	Library was constructed using RNA isolated from breast tissue removed from the left medial lateral breast of a 36-year-old Caucasian female during bilateral simple mastectomy and total breast reconstruction. Pathology indicated benign breast tissue. Patient history included a breast neoplasm, depressive disorder, hyperlipidemia, chronic stomach ulcer, and an ectopic pregnancy. Family history included myocardial infarction, cerebrovascular disease, atherosclerotic coronary artery disease, hyperlipidemia, skin cancer, breast cancer, depressive disorder, esophageal cancer, bone cancer, Hodgkin's lymphoma, bladder cancer, and a heart condition.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
53	SMCCNOS01	Library was constructed using 7.56 X 10 <sup>6</sup> clones from a coronary artery smooth muscle cell library and was subjected to two rounds of subtraction hybridization for 48 hours with 6.12 X 10 <sup>6</sup> clones from a control coronary artery smooth muscle cell library. The starting library for subtraction was constructed using RNA isolated from coronary artery smooth muscle cells removed from a 3-year-old Caucasian male. The cells were treated with TNF alpha & IL-1 beta 10ng/ml each for 20 hours. The hybridization probe for subtraction was derived from a similarly constructed library from RNA isolated from untreated coronary artery smooth muscle cells from the same donor.
54	HUVENOB01	Library was constructed using RNA isolated from HUV-EC-C (ATCC CRL 1730) cells.
55	HNT2RAT01	Library was constructed at Stratagene (STR937231), using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor). Cells were treated with retinoic acid for 24 hours.
56	SINTBST01	Library was constructed using RNA isolated from ileum tissue obtained from an 18-year-old Caucasian female during bowel anastomosis. Pathology indicated Crohn's disease of the ileum, involving 15 cm of the small bowel. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
57	ISLTNOT01	Library was constructed using RNA isolated from a pooled collection of pancreatic islet cells.
58	COLNNOT11	Library was constructed using RNA isolated from colon tissue removed from a 60-year-old Caucasian male during a left hemicolectomy.
59	BONRTUT01	Library was constructed using RNA isolated from rib tumor tissue removed from a 16-year-old Caucasian male during a rib osteotomy and a wedge resection of the lung. Pathology indicated metastatic grade 3 (of 4) osteosarcoma, forming a mass involving the chest wall.
60	LUNGTUT10	Library was constructed using RNA isolated from lung tumor tissue removed from the left upper lobe of a 65-year-old Caucasian female during a segmental lung resection. Pathology indicated metastatic grade 2 myxoid liposarcoma and metastatic grade 4 liposarcoma. Patient history included soft tissue cancer, breast cancer, and secondary lung cancer.
61	OVARNOT09	Library was constructed using RNA isolated from ovarian tissue removed from a 28-year-old Caucasian female during a vaginal hysterectomy and removal of the fallopian tubes and ovaries. Pathology indicated multiple follicular cysts ranging in size from 0.4 to 1.5 cm in the right and left ovaries, chronic cervicitis and squamous metaplasia of the cervix, and endometrium in weakly proliferative phase. Family history included benign hypertension, hyperlipidemia, and atherosclerotic coronary artery disease.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
62	THP1AZS08	Library was constructed using 5.76 million clones from a 5-aza-2'-deoxycytidine (AZ) treated THP-1 promonocyte cell line library. Starting RNA was made from THP-1 promonocyte cells treated for three days with 0.8 micromolar AZ. 5.76 million clones from the AZ-treated THP-1 cell library were then subjected to two rounds of subtractive hybridization with 5 million clones from the untreated THP-1 cell library. Subtractive hybridization conditions were based on the methodologies of Swaroop et al. (1991) Nucleic Acids Res. 19:1954, and Bonaldo et al. (1996) Genome Research 6:791. THP-1 (ATCC TIB 202) is a human promonocyte cell line derived from peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (ref: Int. J. Cancer (1980) 26:171).
63	BRSTNOT12	Library was constructed using RNA isolated from diseased breast tissue removed from a 32-year-old Caucasian female during a bilateral reduction mammoplasty. Pathology indicated nonproliferative fibrocystic disease. Family history included benign hypertension and atherosclerotic coronary artery disease.
64	PENCNOT09	Library was constructed using RNA isolated from penis right corpora cavernosa tissue.
65	PROSTUS19	Library was constructed using 2.36 million clones from a prostate tumor library which was subjected to two rounds of subtraction hybridization with 2.36 million clones from a normal prostate library. The starting library for subtraction was constructed using RNA isolated from prostate tumor tissue removed from a 59-year-old Caucasian male during a radical prostatectomy with regional lymph node excision. Pathology indicated adenocarcinoma (Gleason grade 3+3) involving the prostate peripherally with invasion of the capsule. Adenofibromatous hyperplasia was present. The patient presented with elevated prostate-specific antigen (PSA). Patient history included diverticulitis of the colon, asbestosis, and thrombophlebitis. Family history included benign hypertension, multiple myeloma, hyperlipidemia, and rheumatoid arthritis. The hybridization probe for subtraction was derived from a similarly constructed library, except that NotI-anchored oligo(dT) primer was used. Subtractive hybridization conditions were based on the methodologies of Swaroop et al. (1991) Nucleic Acids Res. 19:1954 and Bonaldo, et al. (1996) Genome Research 6:791.
66	BRABDIR01	Library was constructed using RNA isolated from diseased cerebellum tissue removed from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident. Patient history included Huntington's disease, emphysema, and tobacco abuse.
67	LIVDIR01	Library was constructed using RNA isolated from diseased liver tissue removed from a 63-year-old Caucasian female during a liver transplant. Patient history included primary biliary cirrhosis. Serology was positive for anti-mitochondrial antibody.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
68	COLCDIT03	Library was constructed using RNA isolated from diseased colon polyp tissue removed from the cecum of a 67-year-old female. Pathology indicated a benign cecum polyp. Pathology for the associated tumor tissue indicated invasive grade 3 adenocarcinoma that arose in tubulovillous adenoma forming a fungating mass in the cecum.
69	TBLYNOT01	Library was constructed at Stratagene (STR937214) using RNA isolated from a hybrid of T-B lymphoblasts from an untreated leukemic cell line.
70	KIDNNOT01	Library was constructed using RNA isolated from the kidney tissue of a 64-year-old Caucasian female, who died from an intracranial bleed. Patient history included rheumatoid arthritis and tobacco use.
71	LUNGAST01	Library was constructed using RNA isolated from the lung tissue of a 17-year-old Caucasian male, who died from head trauma. Patient history included asthma.
72	LPARNOT02	Library was constructed using RNA isolated from tissue obtained from the left parotid (salivary) gland of a 70-year-old male with parotid cancer.
73	OVARNOT03	Library was constructed using RNA isolated from ovarian tissue removed from a 43-year-old Caucasian female during a bilateral salpingo-oophorectomy. Pathology for the associated tumor tissue indicated grade 2 mucinous cyst adenocarcinoma. The patient presented with stress incontinence. Patient history included mitral valve disorder, pneumonia, and viral hepatitis. Family history included atherosclerotic coronary artery disease, cerebrovascular disease, breast cancer, and uterine cancer.
74	ENDCNOT03	Library was constructed using RNA isolated from dermal microvascular endothelial cells removed from a neonatal Caucasian male.
75	LUNGNOT18	Library was constructed using RNA isolated from left upper lobe lung tissue removed from a 66-year-old Caucasian female. Pathology for the associated tumor tissue indicated a grade 2 adenocarcinoma. Patient history included cerebrovascular disease, atherosclerotic coronary artery disease, and pulmonary insufficiency. Family history included a myocardial infarction and atherosclerotic coronary artery disease.
76	NGANNOT01	Library was constructed using RNA isolated from tumorous neuroganglion tissue removed from a 9-year-old Caucasian male during a soft tissue excision of the chest wall. Pathology indicated a ganglioneuroma. Family history included asthma.
77	LUNGNOT09	Library was constructed using RNA isolated from lung tumor tissue removed from a 68-year-old Caucasian male during segmental lung resection. Pathology indicated invasive grade 3 squamous cell carcinoma and a metastatic tumor. Patient history included type II diabetes, thyroid disorder, depressive disorder, hyperlipidemia, esophageal ulcer, and tobacco use.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
78	ESOGTUT02	Library was constructed using RNA isolated from esophageal tumor tissue obtained from a 61-year-old Caucasian male during a partial esophagectomy, proximal gastrectomy, pyloromyotomy, and regional lymph node excision. Pathology indicated an invasive grade 3 adenocarcinoma in the esophagus. Family history included atherosclerotic coronary artery disease, type II diabetes, chronic liver disease, primary cardiomyopathy, benign hypertension, and cerebrovascular disease.
79	KIDNNOT19	Library was constructed using RNA isolated from kidney tissue removed a 65-year-old Caucasian male during an exploratory laparotomy and nephroureterectomy. Pathology for the associated tumor tissue indicated a grade 1 renal cell carcinoma within the upper pole of the left kidney. Patient history included malignant melanoma of the abdominal skin, benign neoplasm of colon, cerebrovascular disease, and umbilical hernia. Family history included myocardial infarction, atherosclerotic coronary artery disease, cerebrovascular disease, and prostate cancer.
80	TLYJINT01	Library was constructed using RNA isolated from a Jurkat cell line derived from the T cells of a male. Patient history included acute T-cell leukemia. This is an uninduced Jurkat cell line library from the same donor.
81	PENCNOT06	Library was constructed using RNA isolated from penis corpora cavernosa tissue removed from a 3-year-old Black male. Pathology for the associated tumor tissue indicated invasive grade 4 urothelial carcinoma forming a soft tissue scrotal mass that invaded the cavernous body of the penis and encased both testicles. Right inguinal lymph node showed metastatic grade 4 urothelial carcinoma, with extranodal invasion.
82	UTRSTUT05	Library was constructed using RNA isolated from uterine tumor tissue removed from a 41-year-old Caucasian female during a vaginal hysterectomy with dilation and curettage. Pathology indicated uterine leiomyoma. The endometrium was secretory and contained fragments of endometrial polyps. Benign endo- and ectocervical mucosa were identified in the endocervix. Patient history included a ventral hernia and a benign ovarian neoplasm.
83	LJUNGNOT37	Library was constructed using polyA RNA isolated from lung tissue removed from a 15-year-old Caucasian female who died from a closed head injury. Serology was positive for cytomegalovirus.
84	LIVRTUT09	Library was constructed using RNA isolated from an untreated C3A hepatocytoma cell line which is a derivative of Hep G2, a cell line derived from a hepatoblastoma removed from a 15-year-old Caucasian male.
85	MCLRUNT01	Library was constructed using RNA isolated from untreated peripheral blood mononuclear cell tissue obtained from buffy coat, removed from a 60-year-old male.
86	MCLRUNT01	Library was constructed using RNA isolated from untreated peripheral blood mononuclear cell tissue obtained from buffy coat, removed from a 60-year-old male.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	PE Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	PE Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	PE Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.06E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families



Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score $\geq$ GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPSscan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	



WO 00/78953

PCT/US00/16668

NO:41, SEQ ID NO:42, and SEQ ID NO:43.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

15

5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:51, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, SEQ ID NO:66, SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

30

9. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

35



**WO 00/78953**

**PCT/US00/16668**

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

5

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

10

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

15

17. A pharmaceutical composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, SEQ ID NO:42, and SEQ ID NO:43.

25

18. A method for treating a disease or condition associated with decreased expression of functional TPPT, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 16.

30

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.

35

20. A pharmaceutical composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

WO 00/78953

PCT/US00/16668

21. A method for treating a disease or condition associated with decreased expression of functional TPPT, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 20.

- 5           22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:
- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
  - b) detecting antagonist activity in the sample.

- 10           23. A pharmaceutical composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

- 15           24. A method for treating a disease or condition associated with overexpression of functional TPPT, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 23.

- 25           25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:
- a) combining the polypeptide of claim 1 with at least one test compound under suitable
  - 20 conditions, and
  - b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

- 25           26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:
- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
  - b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound,
  - 30 and
  - c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

- 35           27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method

comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, and
- b) detecting altered expression of the target polynucleotide.

5           28. An isolated polynucleotide comprising a polynucleotide sequence of SEQ ID NO:83.

29. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 28.

10           30. A cell transformed with a recombinant polynucleotide of claim 29.

31. A transgenic organism comprising a recombinant polynucleotide of claim 29.

15           32. A method for producing a polypeptide comprising an amino acid sequence of SEQ ID NO:40, the method comprising:

- a) culturing the cell of claim 30 under conditions suitable for expression of the polypeptide, and
- b) recovering the polypeptide so expressed.

20           33. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 28, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, and
- b) detecting altered expression of the target polynucleotide.

25           34. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

a) an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37.

30           b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37,

c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37, and

35           d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37

35. An isolated polypeptide of claim 34 selected from the group consisting of SEQ ID NO 5, SEQ ID NO:11, and SEQ ID NO:37.

5 36. An isolated polynucleotide encoding a polypeptide of claim 34.

37. An isolated polynucleotide encoding a polypeptide of claim 35

10 38. An isolated polynucleotide of claim 37 selected from the group consisting of SEQ ID NO:48, SEQ ID NO:54, and SEQ ID NO:80.

39. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 36.

15 40. A cell transformed with a recombinant polynucleotide of claim 39.

41. A transgenic organism comprising a recombinant polynucleotide of claim 39.

42. A method for producing a polypeptide of claim 34, the method comprising:

20 a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 34, and

b) recovering the polypeptide so expressed.

25

43. An isolated antibody which specifically binds to a polypeptide of claim 34.

44. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

30 a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:48, SEQ ID NO:54, and SEQ ID NO:80,

b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:48, SEQ ID NO:54, and SEQ ID NO:80.

35 c) a polynucleotide sequence complementary to a),

d) a polynucleotide sequence complementary to b), and

e) an RNA equivalent of a)-d).



45 An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 44.

5 46. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 44, the method comprising:

- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization
- 10 complex is formed between said probe and said target polynucleotide or fragments thereof, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

15 47. A method of claim 46, wherein the probe comprises at least 60 contiguous nucleotides.

48. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 44, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- 20 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof

49. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 34 and a pharmaceutically acceptable excipient.

25 50. A pharmaceutical composition of claim 49, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37.

30 51. A method for treating a disease or condition associated with decreased expression of functional human transport proteins (TPPT), comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 49.

35 52. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 34, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 34 to a compound, and
- b) detecting agonist activity in the sample.

5 54 A method for treating a disease or condition associated with decreased expression of functional human transport proteins (TPPT), comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 53.

10

15

20

34. said method comprising the steps of:

25

claim 34, said method comprising:

30

60. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 38, the method comprising:

- 5           a) exposing a sample comprising the target polynucleotide to a compound, and
- b) detecting altered expression of the target polynucleotide.

61. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:1.

10          62. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:2

63. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:3.

15          64. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:4.

65. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:6.

66. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:7.

20          67. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:8.

68. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:9.

25          69. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:10.

70. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:12.

71. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:13.

30          72. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:14.

73. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:15.

35          74. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:16.

75. A method of claim 9, wherein the polypeptide has the sequence of SEQ ID NO:17.

5

10

15

20

25

30

35

104/6

5

10

15

101. A diagnostic test for a condition or disease associated with the expression of human transport proteins (TPPT) in a biological sample comprising the steps of:

20

a) combining the biological sample with an antibody of claim 10, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and

b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

25

102. A diagnostic test for a condition or disease associated with the expression of human transport proteins (TPPT) in a biological sample comprising the steps of.

a) combining the biological sample with an antibody of claim 43, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and

b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

30

103. The antibody of claim 10, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody.

c) a Fab fragment,

d) a  $F(ab')$ , fragment, or

e) a humanized antibody.

PF-0709 PCT

104. The antibody of claim 43, wherein the antibody is:

- a) a chimeric antibody.
- b) a single chain antibody.
- 5 c) a Fab fragment.
- d) a F(ab')<sub>2</sub> fragment, or
- e) a humanized antibody.

105. A composition comprising an antibody of claim 10 and an acceptable excipient.

106. A composition comprising an antibody of claim 43 and an acceptable excipient.

107. A method of diagnosing a condition or disease associated with the expression of human transport proteins (TPPT) in a subject, comprising administering to said subject an effective amount of the composition of claim 105.

108. A method of diagnosing a condition or disease associated with the expression of human transport proteins (TPPT) in a subject, comprising administering to said subject an effective amount of the composition of claim 106.

109. A composition of claim 105, wherein the antibody is labeled

110. A composition of claim 106, wherein the antibody is labeled

111. A method of diagnosing a condition or disease associated with the expression of human transport proteins (TPPT) in a subject, comprising administering to said subject an effective amount of the composition of claim 109.

112. A method of diagnosing a condition or disease associated with the expression of human transport proteins (TPPT) in a subject, comprising administering to said subject an effective amount of the composition of claim 110

113. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10 comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6.

SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43, or an immunogenic fragment thereof, under conditions to elicit an antibody response.

b) isolating antibodies from said animal; and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43.

114. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 43 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibodies from said animal, and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37.

115. An antibody produced by a method of claim 113.

116. An antibody produced by a method of claim 114.

117. A composition comprising the antibody of claim 115 and a suitable carrier.

118. A composition comprising the antibody of claim 116 and a suitable carrier.

119. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibody producing cells from the animal;

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;

d) culturing the hybridoma cells; and

e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43.

120. A method of making a monoclonal antibody with the specificity of the antibody of claim 43 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibody producing cells from the animal;

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;

d) culturing the hybridoma cells; and

e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37.



5

10

15

20

128. The antibody of claim 43, wherein the antibody is produced by screening a recombinant immunoglobulin library.

30

35

104/11

SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43 in the sample

5

130. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37 in a sample, comprising the steps of:

10 a) incubating the antibody of claim 43 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37 in the sample.

15

131. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43 from a sample, the method comprising:

20

25 a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:41, and SEQ ID NO:43.

30

35 132. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:5, SEQ ID NO:11, and SEQ ID NO:37 from a sample, the method comprising:

a) incubating the antibody of claim 43 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO.5, SEQ ID NO:11, and SEQ ID NO.37.

133 A microarray wherein at least one element of the microarray is a polynucleotide of claim 12.

134. A microarray wherein at least one element of the microarray is a polynucleotide of claim 45.

135 A method for generating a transcript image of a sample which contains polynucleotides, the method comprising the steps of:

a) labeling the polynucleotides of the sample.

b) contacting the elements of the microarray of claim 133 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample.

136. A method for generating a transcript image of a sample which contains polynucleotides, the method comprising the steps of:

a) labeling the polynucleotides of the sample.

b) contacting the elements of the microarray of claim 134 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample

137. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, said target polynucleotide having a sequence of claim 11.

138. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, said target polynucleotide having a sequence of claim 44.

139. An array of claim 137, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide

5 140. An array of claim 138, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

141. An array of claim 139, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

10 142. An array of claim 140, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

143. An array of claim 137, which is a microarray.

15 144. An array of claim 138, which is a microarray.

145. An array of claim 139, further comprising said target polynucleotide hybridized to said first oligonucleotide or polynucleotide.

20 146. An array of claim 140, further comprising said target polynucleotide hybridized to said first oligonucleotide or polynucleotide.

25 147. An array of claim 137, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

148. An array of claim 138, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

30 149. An array of claim 137, wherein each distinct physical location on the substrate contains multiple nucleotide molecules having the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another physical location on the substrate.

35 150. An array of claim 138, wherein each distinct physical location on the substrate contains multiple nucleotide molecules having the same sequence, and each distinct physical location on the

substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another physical location on the substrate.

5 151. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

152 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2

153 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3

10

154. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

155. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

15

156. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

157. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

158. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

20

159 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

160. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

25

161 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

162 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

163. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

30

164 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

165 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

35

166 A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

167. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO 19

168. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20

169. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21

170. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

171. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

172. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

173. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.

174. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

175. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.

176. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

177. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.

178. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

179. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.

180. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.

181. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.

182. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.

183. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.

184. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.

PF-0709 PCT

185. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO.39

186. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO 41

187. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO.43

188. A polypeptide of claim 34, comprising the amino acid sequence of SEQ ID NO.5

189. A polypeptide of claim 34, comprising the amino acid sequence of SEQ ID NO.11

190. A polypeptide of claim 34, comprising the amino acid sequence of SEQ ID NO.37.

191. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:44.

192. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:45.

193. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:46.

194. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:47.

195. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:49.

196. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:50.

197. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:51.

198. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:52.

PF-0709 PCT

199. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:53.

5 200. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:55.

201. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:56

10

202. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:57.

15 203. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:58.

204. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:59.

20 205. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:60

206. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:61.

25

207. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:62

30 208. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:63

209. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:64

35 210. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:65.



212. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:67.

10

15

216. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:71.

20

25

30

222 A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID  
NO:78.

224. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:81.

226. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:84.

15

228. A polynucleotide of claim 44, comprising the polynucleotide sequence of SEQ ID NO.48.

229. A polynucleotide of claim 44, comprising the polynucleotide sequence of SEQ ID  
20 NO:54.

230. A polynucleotide of claim 44, comprising the polynucleotide sequence of SEQ ID NO:80.

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
28 December 2000 (28.12.2000)

PCT

(10) International Publication Number  
**WO 00/78953 A2**

(51) International Patent Classification<sup>7</sup>: C12N 15/12,  
5/10, C07K 14/47, 14/705, 16/18, 16/28, C12Q 1/68,  
A61K 38/17, G01N 33/50, A01K 67/027

(21) International Application Number: PCT/US00/16668

(22) International Filing Date: 16 June 2000 (16.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/139,923 17 June 1999 (17.06.1999) US  
60/148,177 10 August 1999 (10.08.1999) US  
60/149,357 18 August 1999 (18.08.1999) US  
60/162,287 28 October 1999 (28.10.1999) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/139,923 (CIP)
Filed on	17 June 1999 (17.06.1999)
US	60/148,177 (CIP)
Filed on	10 August 1999 (10.08.1999)
US	60/149,357 (CIP)
Filed on	18 August 1999 (18.08.1999)
US	60/162,287 (CIP)
Filed on	28 October 1999 (28.10.1999)

(71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). YANG, Junming [CN/US]; 7125 Bark Lane, San Jose, CA 95129 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). HILLMAN, Jennifer, L.

[US/US]; 230 Monroe Drive, #12, Mountain View, CA 94040 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). BANDMAN, Olga [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). BURFORD, Neil [GB/US]; 1308 4th Avenue, San Francisco, CA 94122 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). AZIMZAI, Yalda [US/US]; 2045 Rocksprings Drive, Hayward, CA 94545 (US). LU, Dyung, Aina, M. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way, #1, Menlo Park, CA 94025 (US).

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: HUMAN TRANSPORT PROTEINS

(57) Abstract: The invention provides human transport proteins (TPPT) and polynucleotides which identify and encode TPPT. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of TPPT.

WO 00/78953 A2

Docket No.: PF-0709 USN

1

# DECLARATION AND POWER OF ATTORNEY FOR UNITED STATES PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

## HUMAN TRANSPORT PROTEINS

the specification of which:

/ X / is attached hereto.

/ was filed on \_\_\_\_\_ as application Serial No. \_\_\_\_\_ and if this box contains an X /, was amended on \_\_\_\_\_.

/ X / was filed as Patent Cooperation Treaty international application No. PCT/US00/16668 on 16 June, 2000, if this box contains an X /, was amended on under Patent Cooperation Treaty Article 19 on \_\_\_\_\_ 2001, and if this box contains an X /, was amended on \_\_\_\_\_.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

**Docket No.: PF-0709 USN**

Country	Number	Filing Date	Priority Claimed
_____	_____	_____	// Yes // No
_____	_____	_____	// Yes // No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
60/139,923	June 17, 1999	Expired
60/148,177	August 10, 1999	Expired
60/149,357	August 18, 1999	Expired
60/162,287	October 28, 1999	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
_____	_____	_____

I hereby appoint the following:

Lucy J. Billings	Reg. No. <u>36,749</u>
Michael C. Cerrone	Reg. No. <u>39,132</u>
Diana Hamlet-Cox	Reg. No. <u>33,302</u>
Richard C. Ekstrom	Reg. No. <u>37,027</u>
Barrie D. Greene	Reg. No. <u>46,740</u>
Lynn E. Murry	Reg. No. <u>42,918</u>
Shirley A. Recipon	Reg. No. <u>47,016</u>
Susan K. Sather	Reg. No. <u>44,316</u>
Michelle M. Stempien	Reg. No. <u>41,327</u>
David G. Streeter	Reg. No. <u>43,168</u>
Stephen Todd	Reg. No. <u>47,139</u>
P. Ben Wang	Reg. No. <u>41,420</u>

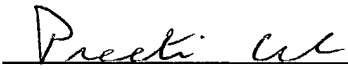
respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

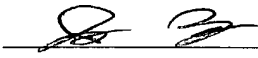
Docket No.: PF-0709 USN

**LEGAL DEPARTMENT**  
**INCYTE GENOMICS, INC.**  
**3160 PORTER DRIVE, PALO ALTO, CA 94304**

**TEL: 650-855-0555      FAX: 650-849-8886 or 650-845-4166**

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

**First Joint Inventor:** 1 - ∞      Full name: Preeti Lal  
 Signature:   
 Date: September 10, 2001  
 Citizenship: India  
 Residence: Santa Clara, California CA  
 P.O. Address: P.O. Box 5142  
Santa Clara, California 95056

**Second Joint Inventor:** 2 - ∞      Full name: Junming Yang  
 Signature:   
 Date: September 17, 2001  
 Citizenship: China  
 Residence: San Jose, California CA  
 P.O. Address: 7125 Bark Lane  
San Jose, California 95129

Docket No.: PF-0709 USN

**Third Joint Inventor:** 3 - ∞ Full name: Henry Yue  
 Signature: Henry Yue.  
 Date: September 24, 2001  
 Citizenship: United States  
 Residence: Sunnyvale, California CA  
 P.O. Address: 826 Lois Avenue  
Sunnyvale, California 94087

**Fourth Joint Inventor:** 4 - ∞ Full name: Jennifer L. Hillman  
 Signature: Jennifer Hillman  
 Date: September 21, 2001  
 Citizenship: United States  
 Residence: Mountain View, California CA  
 P.O. Address: 230 Monroe Drive, #17  
Mountain View, California 94040

**Fifth Joint Inventor:** 5 - ∞ Full name: Y. Tom Tang  
 Signature: Y. Tom Tang  
 Date: Sept. 10, 2001  
 Citizenship: United States  
 Residence: San Jose, California CA  
 P.O. Address: 4230 Ranwick Court  
San Jose, California 95118

Docket No.: PF-0709 USN

**Sixth Joint Inventor:** 6 - 8Full name: Olga BandmanSignature: Olga BandmanDate: 12 September, 2001Citizenship: United StatesResidence: Mountain View, California CAP.O. Address: 366 Anna Avenue  
Mountain View, California 94043**Seventh Joint Inventor:**Full name: Neil Burford

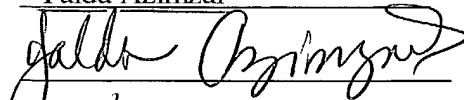
Signature: \_\_\_\_\_


Date: \_\_\_\_\_, 2001


Citizenship: United StatesResidence: Durham, ConnecticutP.O. Address: 105 Wildwood Circle  
Durham, Connecticut 06422**Eighth Joint Inventor:** 8 - 8Full name: Mariah R. BaughnSignature: M. R. BaughnDate: September 5, 2001Citizenship: United StatesResidence: San Leandro, California CAP.O. Address: 14244 Santiago Road  
San Leandro, California 94577



Docket No.: PF-0709 USN

**Ninth Joint Inventor:** 9 - ∞ Full name: Yalda Azimzai  
 Signature:   
 Date: September 13, 2001  
 Citizenship: United States  
 Residence: Castro Valley, California CA  
 P.O. Address: 5518 Boulder Canyon Drive  
Castro Valley, California 94552

**Tenth Joint Inventor:** 10 - ∞ Full name: Dyung Aina M. Lu  
 Signature:   
 Date: Sept 7, 2001  
 Citizenship: United States  
 Residence: San Jose, California CA  
 P.O. Address: 233 Coy Drive  
San Jose, California 95123

**Eleventh Joint Inventor:** 11 - ∞ Full name: Janice Au-Young  
 Signature:   
 Date: September 7, 2001  
 Citizenship: United States  
 Residence: Brisbane, California CA  
 P.O. Address: 233 Golden Eagle Lane  
Brisbane, California 94005

Docket No.: PF-0709 USN

Twelfth Joint Inventor: 12-00 Full name:

Chandra PattersonArvizu <sup>CA</sup> 9/10/01

Signature:

Chandra Arvizu

Date:

September 10, 2001

Citizenship:

United States

Residence:

Menlo Park, California CA

P.O. Address:

490 Sherwood Way, #1

Menlo Park, California 94025

Docket No.: PF-0709 USN

**DECLARATION AND POWER OF ATTORNEY FOR  
UNITED STATES PATENT APPLICATION**

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name, and

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if more than one name is listed below) of the subject matter which is claimed and for which a United States patent is sought on the invention entitled

**HUMAN TRANSPORT PROTEINS**

the specification of which:

/ X / is attached hereto.

/ / was filed on \_\_\_\_\_ as application Serial No. \_\_\_\_\_ and if this box contains an X / /, was amended on \_\_\_\_\_.

/ X / was filed as Patent Cooperation Treaty international application No. PCT/US00/16668 on 16 June, 2000, if this box contains an X / /, was amended on under Patent Cooperation Treaty Article 19 on \_\_\_\_\_ 2001, and if this box contains an X / /, was amended on \_\_\_\_\_.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge my duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim the benefit under Title 35, United States Code, §119 or §365(a)-(b) of any foreign application(s) for patent or inventor's certificate indicated below and of any Patent Cooperation Treaty international applications(s) designating at least one country other than the United States indicated below and have also identified below any foreign application(s) for patent or inventor's certificate and Patent Cooperation Treaty international application(s) designating at least one country other than the United States for the same subject matter and having a filing date before that of the application for said subject matter the priority of which is claimed:

**Docket No.: PF-0709 USN**

Country	Number	Filing Date	Priority Claimed
_____	_____	_____	// Yes // No
_____	_____	_____	// Yes // No

I hereby claim the benefit under Title 35, United States Code, §119(e) of any United States provisional application(s) listed below.

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
60/139,923	June 17, 1999	Expired
60/148,177	August 10, 1999	Expired
60/149,357	August 18, 1999	Expired
60/162,287	October 28, 1999	Expired

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in said prior application(s) in the manner required by the first paragraph of Title 35, United States Code §112, I acknowledge my duty to disclose material information as defined in Title 37 Code of Federal Regulations, §1.56(a) which occurred between the filing date(s) of the prior application(s) and the national or Patent Cooperation Treaty international filing date of this application:

Application Serial No.	Filed	Status (Pending, Abandoned, Patented)
_____	_____	_____

I hereby appoint the following:

Lucy J. Billings	Reg. No. 36,749
Michael C. Cerrone	Reg. No. 39,132
Diana Hamlet-Cox	Reg. No. 33,302
Richard C. Ekstrom	Reg. No. 37,027
Barrie D. Greene	Reg. No. 46,740
Lynn E. Murry	Reg. No. 42,918
Shirley A. Recipon	Reg. No. 47,016
Susan K. Sather	Reg. No. 44,316
Michelle M. Stempien	Reg. No. 41,327
David G. Streeter	Reg. No. 43,168
Stephen Todd	Reg. No. 47,139
P. Ben Wang	Reg. No. 41,420

respectively and individually, as my patent attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith. Please address all communications to:

Docket No.: PF-0709 USN

**LEGAL DEPARTMENT  
INCYTE GENOMICS, INC.  
3160 PORTER DRIVE, PALO ALTO, CA 94304**

**TEL: 650-855-0555      FAX: 650-849-8886 or 650-845-4166**

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

<b>First Joint Inventor:</b>	Full name:	<u>Preeti Lal</u>
	Signature:	_____
	Date:	_____, 2001
	Citizenship:	<u>India</u>
	Residence:	<u>Santa Clara, California</u>
	P.O. Address:	<u>P.O. Box 5142</u> <u>Santa Clara, California 95056</u>
<b>Second Joint Inventor:</b>	Full name:	<u>Junming Yang</u>
	Signature:	_____
	Date:	_____, 2001
	Citizenship:	<u>China</u>
	Residence:	<u>San Jose, California</u>
	P.O. Address:	<u>7125 Bark Lane</u> <u>San Jose, California 95129</u>

Docket No.: PF-0709 USN

**Third Joint Inventor:**

Full name: Henry Yue  
Signature: \_\_\_\_\_  
Date: \_\_\_\_\_, 2001  
Citizenship: United States  
Residence: Sunnyvale, California  
P.O. Address: 826 Lois Avenue  
Sunnyvale, California 94087

**Fourth Joint Inventor:**

Full name: Jennifer L. Hillman  
Signature: \_\_\_\_\_  
Date: \_\_\_\_\_, 2001  
Citizenship: United States  
Residence: Mountain View, California  
P.O. Address: 230 Monroe Drive, #17  
Mountain View, California 94040

**Fifth Joint Inventor:**

Full name: Y. Tom Tang  
Signature: \_\_\_\_\_  
Date: \_\_\_\_\_, 2001  
Citizenship: United States  
Residence: San Jose, California  
P.O. Address: 4230 Ranwick Court  
San Jose, California 95118

**Docket No.: PF-0709 USN****Sixth Joint Inventor:**

Full name: Olga Bandman

Signature: \_\_\_\_\_

Date: \_\_\_\_\_, 2001

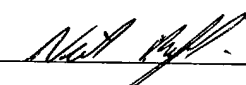
Citizenship: United States

Residence: Mountain View, California

P.O. Address: 366 Anna Avenue  
Mountain View, California 94043

**Seventh Joint Inventor: G -- ∞**

Full name: Neil Burford

Signature: 

Date: 9/21, 2001

Citizenship: 9/21/01  
NP. United States UNITED KINGDOM

Residence: Durham, Connecticut CT

P.O. Address: 105 Wildwood Circle  
Durham, Connecticut 06422

**Eighth Joint Inventor:**

Full name: Mariah R. Baughn

Signature: \_\_\_\_\_

Date: \_\_\_\_\_, 2001

Citizenship: United States

Residence: San Leandro, California

P.O. Address: 14244 Santiago Road  
San Leandro, California 94577

**Docket No.: PF-0709 USN**

**Ninth Joint Inventor:**

Full name: Yalda Azimzai

Signature: \_\_\_\_\_

Date: \_\_\_\_\_, 2001

Citizenship: United States

Residence: Castro Valley, California

P.O. Address: 5518 Boulder Canyon Drive  
Castro Valley, California 94552

**Tenth Joint Inventor:**

Full name: Dyung Aina M. Lu

Signature: \_\_\_\_\_

Date: \_\_\_\_\_, 2001

Citizenship: United States

Residence: San Jose, California

P.O. Address: 233 Coy Drive  
San Jose, California 95123

**Eleventh Joint Inventor:**

Full name: Janice Au-Young

Signature: \_\_\_\_\_

Date: \_\_\_\_\_, 2001

Citizenship: United States

Residence: Brisbane, California

P.O. Address: 233 Golden Eagle Lane  
Brisbane, California 94005



**Docket No.: PF-0709 USN**

**Twelfth Joint Inventor:**

Full name: Chandra Patterson

Signature: \_\_\_\_\_

Date: \_\_\_\_\_, 2001

Citizenship: United States

Residence: Menlo Park, California

P.O. Address: 490 Sherwood Way, #1  
Menlo Park, California 94025

WO 00/78953

PCT/US00/16668

SEQUENCE LISTING

<110> INCYTE GENOMICS, INC.  
LAL, Preeti  
YANG, Junming  
YUE, Henry  
HILLMAN, Jennifer L.  
TANG, Y. Tom  
BANDMAN, Olga  
BURFORD, Neil  
BAUGHN, Mariah R.  
AZIMZAI, Yalda  
LU, Dyung Aina M.  
AU-YOUNG, Janice  
PATTERSON, Chandra

<120> HUMAN TRANSPORT PROTEINS

<130> PF-0709 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/139,923; 60/148,177; 60/149,357; 60/162,287

<151> 1999-06-17; 1999-08-10; 1999-08-18; 1999-10-28

<160> 86

<170> PERL Program

<210> 1

<211> 623

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 264114CD1

<400> 1

Met	Ser	Thr	Gln	Asp	Glu	Arg	Gln	Ile	Asn	Thr	Glu	Tyr	Ala	Val
1				5					10					15
Ser	Leu	Leu	Glu	Gln	Leu	Lys	Leu	Phe	Tyr	Glu	Gln	Gln	Leu	Phe
				20					25					30
Thr	Asp	Ile	Val	Leu	Ile	Val	Glu	Gly	Thr	Glu	Phe	Pro	Cys	His
				35					40					45
Lys	Met	Val	Leu	Ala	Thr	Cys	Ser	Ser	Tyr	Phe	Arg	Ala	Met	Phe
				50					55					60
Met	Ser	Gly	Leu	Ser	Glu	Ser	Lys	Gln	Thr	His	Val	His	Leu	Arg
				65					70					75
Asn	Val	Asp	Ala	Ala	Thr	Leu	Gln	Ile	Ile	Ile	Thr	Tyr	Ala	Tyr
				80					85					90
Thr	Gly	Asn	Leu	Ala	Met	Asn	Asp	Ser	Thr	Val	Glu	Gln	Leu	Tyr
				95					100					105
Glu	Thr	Ala	Cys	Phe	Leu	Gln	Val	Glu	Asp	Val	Leu	Gln	Arg	Cys
				110					115					120
Arg	Glu	Tyr	Leu	Ile	Lys	Lys	Ile	Asn	Ala	Glu	Asn	Cys	Val	Arg
				125					130					135
Leu	Leu	Ser	Phe	Ala	Asp	Leu	Phe	Ser	Cys	Glu	Glu	Leu	Lys	Gln
				140					145					150
Ser	Ala	Lys	Arg	Met	Val	Glu	His	Lys	Phe	Thr	Ala	Val	Tyr	His
				155					160					165
Gln	Asp	Ala	Phe	Met	Gln	Leu	Ser	His	Asp	Leu	Leu	Ile	Asp	Ile
				170					175					180
Leu	Ser	Ser	Asp	Asn	Leu	Asn	Val	Glu	Lys	Glu	Glu	Thr	Val	Arg
				185					190					195
Glu	Ala	Ala	Met	Leu	Trp	Leu	Glu	Tyr	Asn	Thr	Glu	Ser	Arg	Ser

WO 00/78953

PCT/US00/16668

Gln Tyr Leu Ser	200	Ser Val Leu Ser Gln	205	Ile Arg Ile Asp Ala	210
Ser Glu Val Thr	215	Gln Arg Ala Trp Phe	220	Gly Leu Pro Pro	225
Asp Lys Ser Val	230	Val Val Gln Gly Leu	235	Tyr Lys Ser Met Pro	240
Phe Phe Lys Pro	245	Leu Gly Met Thr	250	Glu Glu Met Met	255
Phe Ile Glu Ala	260	Ser Ser Glu Asn Pro	265	Cys Ser Leu Tyr Ser	270
Val Cys Tyr Ser	275	Pro Gln Ala Glu Lys	280	Val Tyr Lys Leu Cys	285
Pro Pro Ala Asp	290	Leu His Lys Val Gly	295	Thr Val Val Thr Pro	300
Asn Asp Ile Tyr	305	Ile Ala Gly Gly Gln	310	Val Pro Leu Lys Asn	315
Lys Thr Asn His	320	Ser Lys Thr Ser Lys	325	Leu Gln Thr Ala Phe	330
Thr Val Asn Cys	335	Phe Tyr Trp Phe Asp	340	Ala Gln Gln Asn Thr	345
Phe Pro Lys Thr	350	Pro Met Leu Phe Val	355	Arg Ile Lys Pro Ser	360
Val Cys Cys Glu	365	Gly Tyr Ile Tyr Ala	370	Ile Gly Gly Asp Ser	375
Gly Gly Glu Leu	380	Asn Arg Arg Thr Val	385	Glu Arg Tyr Asp Thr	390
Lys Asp Glu Trp	395	Thr Met Val Ser Pro	400	Leu Pro Cys Ala Trp	405
Trp Ser Ala Ala	410	Val Val Val His Asp	415	Cys Ile Tyr Val Met	420
Leu Asn Leu Met	425	Tyr Cys Tyr Phe Pro	430	Arg Ser Asp Ser Trp	435
Glu Met Ala Met	440	Arg Gln Thr Ser Arg	445	Ser Phe Ala Ser Ala	450
Ala Phe Gly Asp	455	Ile Phe Tyr Ile Gly	460	Gly Leu His Ile Ala	465
Thr Asn Ser Gly	470	Ile Arg Leu Pro Ser	475	Gly Thr Val Asp Gly	480
Ser Val Thr Val	485	Glu Ile Tyr Asp Val	490	Asn Lys Asn Glu Trp	495
Met Ala Ala Asn	500	Ile Pro Ala Lys Arg	505	Tyr Ser Asp Pro Cys	510
Arg Ala Val Val	515	Ile Ser Asn Ser Leu	520	Cys Val Phe Met Arg	525
Thr His Leu Asn	530	Glu Arg Ala Lys Tyr	535	Val Thr Tyr Gln Tyr	540
Leu Glu Leu Asp	545	Arg Trp Ser Leu Arg	550	Gln His Ile Ser Glu	555
Val Leu Trp Asp	560	Leu Gly Arg Asp Phe	565	Arg Cys Thr Val Gly	570
Leu Tyr Pro Ser	575	Cys Leu Glu Glu Ser	580	Pro Trp Lys Pro Pro	585
Tyr Leu Phe Ser	590	Thr Asp Gly Thr Glu	595	Glu Phe Glu Leu Asp	600
Glu Met Val Ala	605	Leu Pro Pro Val	610		615
	620				

<210> 2  
 <211> 99  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1455669CD1

<400> 2  
 Met Ala Ala Pro Ala Glu Pro Cys Ala Gly Gln Gly Val Trp Asn

WO 00/78953

PCT/US00/16668

1	5	10	15
Gln Thr Glu Pro	Glu Pro Ala Ala Thr	Ser Leu Leu Ser Leu	Cys
	20	25	30
Phe Leu Arg Thr	Ala Gly Val Trp Val	Pro Pro Met Tyr Leu	Trp
	35	40	45
Val Leu Gly Pro	Ile Tyr Leu Leu Phe	Ile His His His Gly	Arg
	50	55	60
Gly Tyr Leu Arg	Met Ser Pro Leu Phe	Lys Ala Lys Met Val	Ala
	65	70	75
Ala Ile Pro Gly	Ser Leu Glu Pro Gly	Asn Val Arg Gly Arg	Gln
	80	85	90
Gly Thr Gly Trp	Asn Leu Val Lys Ser		
	95		

<210> 3

<211> 374

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2084989CD1

<400> 3

Met Glu Ser Lys	Met Gly Glu Leu Pro	Leu Asp Ile Asn Ile	Gln
1	5	10	15
Glu Pro Arg Trp	Asp Gln Ser Thr Phe	Leu Gly Arg Ala Arg	His
	20	25	30
Phe Phe Thr Val	Thr Asp Pro Arg Asn	Leu Leu Ser Gly	Ala
	35	40	45
Gln Leu Glu Ala	Ser Arg Asn Ile Val	Gln Asn Tyr Arg	Ala
	50	55	60
Val Val Thr Pro	Gly Ile Thr Glu Asp	Gln Leu Trp Arg	Ala
	65	70	75
Tyr Val Tyr Asp	Ser Ala Phe His Pro	Asp Thr Gly Glu Lys	Val
	80	85	90
Val Leu Ile Gly	Arg Met Ser Ala Gln	Val Pro Met Asn Met	Thr
	95	100	105
Ile Thr Gly Cys	Met Leu Thr Phe Tyr	Arg Lys Thr Pro Thr	Val
	110	115	120
Val Phe Trp Gln	Trp Val Asn Gln Ser	Phe Asn Ala Ile Val	Asn
	125	130	135
Tyr Ser Asn Arg	Ser Gly Asp Thr Pro	Ile Thr Val Arg Gln	Leu
	140	145	150
Gly Thr Ala Tyr	Val Ser Ala Thr Thr	Gly Ala Val Ala Thr	Ala
	155	160	165
Leu Gly Leu Lys	Ser Leu Thr Lys His	Leu Pro Pro Leu Val	Gly
	170	175	180
Arg Phe Val Pro	Phe Ala Ala Val Ala	Ala Ala Asn Cys Ile	Asn
	185	190	195
Ile Pro Leu Met	Arg Gln Arg Glu Leu	Gln Val Gly Ile Pro	Val
	200	205	210
Ala Asp Glu Ala	Gly Gln Arg Leu Gly	Tyr Ser Val Thr Ala	Ala
	215	220	225
Lys Gln Gly Ile	Phe Gln Val Val Ile	Ser Arg Ile Cys Met	Ala
	230	235	240
Ile Pro Ala Met	Ala Ile Pro Pro Leu	Ile Met Asp Thr Leu	Glu
	245	250	255
Lys Lys Asp Phe	Leu Lys Val Gly Asp	Cys Thr Ser Leu Val	Leu
	260	265	270
Glu Trp Ala Met	Ala Gly Arg Ser Asp	Gln Ala Pro Thr Leu	Ser
	275	280	285
Pro Ala Ser Pro	Asp Ser Leu Arg Leu	Ala Ser Pro Ser Pro	Asp
	290	295	300
Pro Cys Thr Ala	Ser Ser Thr Phe Val	His Ser Ala Arg Met	Asn
	305	310	315
Trp Ala Gly Val	Lys Glu Leu Cys Arg	Gly Arg Arg Arg Gly	Gln
	320	325	330
Arg Lys Glu Thr	Asn Phe Ile Ser Val	Thr Pro Val Ala Ser	Asp





WO 00/78953

PCT/US00/16668

	140		145		150
Glu Ile Pro Phe	Ser Leu Val Gln Phe	Pro Leu Trp Glu Ser	Leu		
	155		160		165
Lys Ala Leu Trp	Ser Trp Arg Gln Asp	His Val Val Asp Ser	Trp		
	170		175		180
Gln Ser Ala Val	Cys Gly Ala Phe Ala	Gly Gly Phe Ala Ala	Ala		
	185		190		195
Val Thr Thr Pro	Leu Asp Val Ala Lys	Thr Arg Ile Thr Leu	Ala		
	200		205		210
Lys Ala Gly Ser	Ser Thr Ala Asp Gly	Asn Val Leu Ser Val	Leu		
	215		220		225
His Gly Val Trp	Arg Ser Gln Gly Leu	Ala Gly Leu Phe Ala	Gly		
	230		235		240
Val Phe Pro Arg	Met Ala Ala Ile Ser	Leu Gly Gly Phe Ile	Phe		
	245		250		255
Leu Gly Ala Tyr	Asp Arg Thr His Ser	Leu Leu Glu Val	Gly		
	260		265		270
Arg Lys Ser Pro					

<210> 7  
 <211> 291  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 876677CD1

<400> 7

Met Asp Ser Arg Val	Ser Ser Pro Glu Lys	Gln Asp Lys Glu Asn
1	5	10
Phe Val Gly Val	Asn Asn Lys Arg Leu	Gly Val Cys Gly Trp Ile
	20	25
Leu Phe Ser Leu	Ser Phe Leu Leu Val	Ile Ile Thr Phe Pro Ile
	35	40
Ser Ile Trp Met	Cys Leu Lys Ile Ile	Lys Glu Tyr Glu Arg Ala
	50	55
Val Val Phe Arg	Leu Gly Arg Ile Gln	Ala Asp Lys Ala Lys Gly
	65	70
Pro Gly Leu Ile	Leu Val Leu Pro Cys	Ile Asp Val Phe Val Lys
	80	85
Val Asp Leu Arg	Thr Val Thr Cys Asn	Ile Pro Pro Gln Glu Ile
	95	100
Leu Thr Arg Asp	Ser Val Thr Thr Gln	Val Asp Gly Val Val Tyr
	110	115
Tyr Arg Ile Tyr	Ser Ala Val Ser Ala	Val Ala Asn Val Asn Asp
	125	130
Val His Gln Ala	Thr Phe Leu Leu Ala	Gln Thr Thr Leu Arg Asn
	140	145
Val Leu Gly Thr	Gln Thr Leu Ser Gln	Ile Leu Ala Gly Arg Glu
	155	160
Glu Ile Ala His	Ser Ile Gln Thr Leu	Leu Asp Asp Ala Thr Glu
	170	175
Leu Trp Gly Ile	Arg Val Ala Arg Val	Glu Ile Lys Asp Val Arg
	185	190
Ile Pro Val Gln	Leu Gln Arg Ser Met	Ala Ala Glu Ala Glu Ala
	200	205
Thr Arg Glu Ala	Arg Ala Lys Val Leu	Ala Ala Glu Gly Glu Met
	215	220
Asn Ala Ser Lys	Ser Leu Lys Ser Ala	Ser Met Val Leu Ala Glu
	230	235
Ser Pro Ile Ala	Leu Gln Leu Arg Tyr	Leu Gln Thr Leu Ser Thr
	245	250
Val Ala Thr Glu	Lys Asn Ser Thr Ile	Val Phe Pro Leu Pro Met
	260	265
Asn Ile Leu Glu	Gly Ile Gly Gly Val	Ser Tyr Asp Asn His Lys
	275	280
Lys Leu Pro Asn	Lys Ala	

290

<210> 8  
 <211> 381  
 <212> PRT  
 <213> Homo sapiens  
 <220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2326143CD1

<400> 8  
 Met Ser Arg His Glu Gly Val Ser Cys Asp Ala Cys Leu Lys Gly  
 1 5 10 15  
 Asn Phe Arg Gly Arg Arg Tyr Lys Cys Leu Ile Cys Tyr Asp Tyr  
 20 25 30  
 Asp Leu Cys Ala Ser Cys Tyr Glu Ser Gly Ala Thr Thr Thr Arg  
 35 40 45  
 His Thr Thr Asp His Pro Met Gln Cys Ile Leu Thr Arg Val Asp  
 50 55 60  
 Phe Asp Leu Tyr Tyr Gly Gly Glu Ala Phe Ser Val Glu Gln Pro  
 65 70 75  
 Gln Ser Phe Thr Cys Pro Tyr Cys Gly Lys Met Gly Tyr Thr Glu  
 80 85 90  
 Thr Ser Leu Gln Glu His Val Thr Ser Glu His Ala Glu Thr Ser  
 95 100 105  
 Thr Glu Val Ile Cys Pro Ile Cys Ala Ala Leu Pro Gly Gly Asp  
 110 115 120  
 Pro Asn His Val Thr Asp Asp Phe Ala Ala His Leu Thr Leu Glu  
 125 130 135  
 His Arg Ala Pro Arg Asp Leu Asp Glu Ser Ser Gly Val Arg His  
 140 145 150  
 Val Arg Arg Met Phe His Pro Gly Arg Gly Leu Gly Gly Pro Arg  
 155 160 165  
 Ala Arg Arg Ser Asn Met His Phe Thr Ser Ser Ser Thr Gly Gly  
 170 175 180  
 Leu Ser Ser Ser Gln Ser Ser Tyr Ser Pro Ser Asn Arg Glu Ala  
 185 190 195  
 Met Asp Pro Ile Ala Glu Leu Leu Ser Gln Leu Ser Gly Val Arg  
 200 205 210  
 Arg Ser Ala Gly Gly Gln Leu Asn Ser Ser Gly Pro Ser Ala Ser  
 215 220 225  
 Gln Leu Gln Gln Leu Gln Met Gln Leu Gln Leu Glu Arg Gln His  
 230 235 240  
 Ala Gln Ala Ala Arg Gln Gln Leu Glu Thr Ala Arg Asn Ala Thr  
 245 250 255  
 Arg Arg Thr Asn Thr Ser Ser Val Thr Thr Thr Ile Thr Gln Ser  
 260 265 270  
 Thr Ala Thr Thr Asn Ile Ala Asn Thr Glu Ser Ser Gln Gln Thr  
 275 280 285  
 Leu Gln Asn Ser Gln Phe Leu Leu Thr Arg Leu Asn Asp Pro Lys  
 290 295 300  
 Met Ser Glu Thr Glu Arg Gln Ser Met Glu Ser Glu Arg Ala Asp  
 305 310 315  
 Arg Ser Leu Phe Val Gln Glu Leu Leu Leu Ser Thr Leu Val Arg  
 320 325 330  
 Glu Glu Ser Ser Ser Ser Asp Glu Asp Asp Arg Gly Glu Met Ala  
 335 340 345  
 Asp Phe Gly Ala Met Gly Cys Val Asp Ile Met Pro Leu Asp Val  
 350 355 360  
 Ala Leu Glu Asn Leu Asn Leu Lys Glu Ser Asn Lys Gly Asn Glu  
 365 370 375  
 Pro Pro Pro Pro Pro Leu  
 380

<210> 9  
 <211> 190  
 <212> PRT  
 <213> Homo sapiens







WO 00/78953

PCT/US00/16668

&lt;223&gt; Incyte ID No: 1429651CD1

&lt;400&gt; 13

```

Met Glu Ser Gly Thr Ser Ser Pro Gln Pro Pro Gln Leu Asp Pro
 1      5      10      15
Leu Asp Ala Phe Pro Gln Lys Gly Leu Glu Pro Gly Asp Ile Ala
 20      25      30
Val Leu Val Leu Tyr Phe Leu Phe Val Leu Ala Val Gly Leu Trp
 35      40      45
Ser Thr Val Lys Thr Lys Arg Asp Thr Val Lys Gly Tyr Phe Leu
 50      55      60
Ala Gly Gly Asp Met Val Trp Trp Pro Val Gly Ala Ser Leu Phe
 65      70      75
Ala Ser Asn Val Gly Ser Gly His Phe Ile Gly Leu Ala Gly Ser
 80      85      90
Gly Ala Ala Thr Gly Ile Ser Val Ser Ala Tyr Glu Leu Asn Gly
 95      100      105
Leu Phe Ser Val Leu Met Leu Ala Trp Ile Phe Leu Pro Ile Tyr
110      115      120
Ile Ala Gly Gln Val Thr Thr Met Pro Glu Tyr Leu Arg Lys Arg
125      130      135
Phe Gly Gly Ile Arg Ile Pro Ile Ile Leu Ala Val Leu Tyr Leu
140      145      150
Phe Ile Tyr Ile Phe Thr Lys Ile Ser Val Asp Met Tyr Ala Gly
155      160      165
Ala Ile Phe Ile Gln Gln Ser Leu His Leu Asp Leu Tyr Leu Ala
170      175      180
Ile Val Gly Leu Leu Ala Ile Thr Ala Val Tyr Thr Val Ala Gly
185      190      195
Gly Leu Ala Ala Val Ile Tyr Thr Asp Ala Leu Gln Thr Leu Ile
200      205      210
Met Leu Ile Gly Ala Leu Thr Leu Met Gly Tyr Ser Phe Ala Ala
215      220      225
Val Gly Gly Met Glu Gly Leu Lys Glu Lys Tyr Phe Leu Ala Leu
230      235      240
Ala Ser Asn Arg Ser Glu Asn Ser Ser Cys Gly Leu Pro Arg Glu
245      250      255
Asp Ala Phe His Ile Phe Arg Asp Pro Leu Thr Ser Asp Leu Pro
260      265      270
Trp Pro Gly Val Leu Phe Gly Met Ser Ile Pro Ser Leu Trp Tyr
275      280      285
Trp Cys Thr Asp Gln Val Ile Val Gln Arg Thr Leu Ala Ala Lys
290      295      300
Asn Leu Ser His Ala Lys Gly Gly Ala Leu Met Ala Ala Tyr Leu
305      310      315
Lys Val Leu Pro Leu Phe Ile Met Val Phe Pro Gly Met Val Ser
320      325      330
Arg Ile Leu Phe Pro Asp Gln Val Ala Cys Ala Asp Pro Glu Ile
335      340      345
Cys Gln Lys Ile Cys Ser Asn Pro Ser Gly Cys Ser Asp Ile Ala
350      355      360
Tyr Pro Lys Leu Val Leu Glu Leu Leu Pro Thr Gly Leu Arg Gly
365      370      375
Leu Met Met Ala Val Met Val Ala Ala Leu Met Ser Ser Leu Thr
380      385      390
Ser Ile Phe Asn Ser Ala Ser Thr Ile Phe Thr Met Asp Leu Trp
395      400      405
Asn His Leu Arg Pro Arg Ala Ser Glu Lys Glu Leu Met Ile Val
410      415      420
Gly Arg Val Phe Val Leu Leu Leu Val Leu Val Ser Ile Leu Trp
425      430      435
Ile Pro Val Val Gln Ala Ser Gln Gly Gly Gln Leu Phe Ile Tyr
440      445      450
Ile Gln Ser Ile Ser Ser Tyr Leu Gln Pro Pro Val Ala Val Val
455      460      465
Phe Ile Met Gly Cys Phe Trp Lys Arg Thr Asn Glu Lys Gly Ala
470      475      480
Phe Trp Gly Leu Ile Ser Gly Leu Leu Leu Gly Leu Val Arg Leu

```

WO 00/78953

PCT/US00/16668

Val	Leu	Asp	Phe	Ile	Tyr	Val	Gln	Pro	Arg	Cys	Asp	Gln	Pro	Asp
485									490					495
500									505					510
Glu	Arg	Pro	Val	Leu	Val	Lys	Ser	Ile	His	Tyr	Leu	Tyr	Phe	Ser
515									520					525
Met	Ile	Leu	Ser	Thr	Val	Thr	Leu	Ile	Thr	Val	Ser	Thr	Val	Ser
530									535					540
Trp	Phe	Thr	Glu	Pro	Pro	Ser	Lys	Glu	Met	Val	Ser	His	Leu	Thr
545									550					555
Trp	Phe	Thr	Arg	His	Asp	Pro	Val	Val	Gln	Lys	Glu	Gln	Ala	Pro
560									565					570
Pro	Ala	Ala	Pro	Leu	Ser	Leu	Thr	Leu	Ser	Gln	Asn	Gly	Met	Pro
575									580					585
Glu	Ala	Ser	Ser	Ser	Ser	Ser	Val	Gln	Phe	Glu	Met	Val	Gln	Glu
590									595					600
Asn	Thr	Ser	Lys	Thr	His	Ser	Cys	Asp	Met	Thr	Pro	Lys	Gln	Ser
605									610					615
Lys	Val	Val	Lys	Ala	Ile	Leu	Trp	Leu	Cys	Gly	Ile	Gln	Glu	Lys
620									625					630
Gly	Lys	Glu	Glu	Leu	Pro	Ala	Arg	Ala	Glu	Ala	Ile	Ile	Val	Ser
635									640					645
Leu	Glu	Glu	Asn	Pro	Leu	Val	Lys	Thr	Leu	Leu	Asp	Val	Asn	Leu
650									655					660
Ile	Phe	Cys	Val	Ser	Cys	Ala	Ile	Phe	Ile	Trp	Gly	Tyr	Phe	Ala
665									670					675

<210> 14  
 <211> 320  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2069971CD1

<400> 14

Met	Tyr	His	Cys	His	Ser	Gly	Ser	Lys	Pro	Thr	Glu	Lys	Gly	Ala
1				5					10					15
Asn	Glu	Tyr	Ala	Tyr	Ala	Lys	Trp	Lys	Leu	Cys	Ser	Ala	Ser	Ala
				20					25					30
Ile	Cys	Phe	Ile	Phe	Met	Ile	Ala	Glu	Val	Val	Gly	Gly	His	Ile
				35					40					45
Ala	Gly	Ser	Leu	Ala	Val	Val	Thr	Asp	Ala	Ala	His	Leu	Leu	Ile
				50					55					60
Asp	Leu	Thr	Ser	Phe	Leu	Leu	Ser	Leu	Phe	Ser	Leu	Trp	Leu	Ser
				65					70					75
Ser	Lys	Pro	Pro	Ser	Lys	Arg	Leu	Thr	Phe	Gly	Trp	His	Arg	Ala
				80					85					90
Glu	Ile	Leu	Gly	Ala	Leu	Leu	Ser	Ile	Leu	Cys	Ile	Trp	Val	Val
				95					100					105
Thr	Gly	Val	Leu	Val	Tyr	Leu	Ala	Cys	Glu	Arg	Leu	Leu	Tyr	Pro
				110					115					120
Asp	Tyr	Gln	Ile	Gln	Ala	Thr	Val	Met	Ile	Ile	Val	Ser	Ser	Cys
				125					130					135
Ala	Val	Ala	Ala	Asn	Ile	Val	Leu	Thr	Val	Val	Leu	His	Gln	Arg
				140					145					150
Cys	Leu	Gly	His	Asn	His	Lys	Glu	Val	Gln	Ala	Asn	Ala	Ser	Val
				155					160					165
Arg	Ala	Ala	Phe	Val	His	Ala	Leu	Gly	Asp	Leu	Phe	Gln	Ser	Ile
				170					175					180
Ser	Val	Leu	Ile	Ser	Ala	Leu	Ile	Ile	Tyr	Phe	Lys	Pro	Glu	Tyr
				185					190					195
Lys	Ile	Ala	Asp	Pro	Ile	Cys	Thr	Phe	Ile	Phe	Ser	Ile	Leu	Val
				200					205					210
Leu	Ala	Ser	Thr	Ile	Thr	Ile	Leu	Lys	Asp	Phe	Ser	Ile	Leu	Leu
				215					220					225
Met	Glu	Gly	Val	Pro	Lys	Ser	Leu	Asn	Tyr	Ser	Gly	Val	Lys	Glu
				230					235					240

WO 00/78953

PCT/US00/16668

Leu	Ile	Leu	Ala	Val	Asp	Gly	Val	Leu	Ser	Val	His	Ser	Leu	His
				245					250					255
Ile	Trp	Ser	Leu	Thr	Met	Asn	Gln	Val	Ile	Leu	Ser	Ala	His	Val
				260					265					270
Ala	Thr	Ala	Ala	Ser	Arg	Asp	Ser	Gln	Val	Val	Arg	Arg	Glu	Ile
				275					280					285
Ala	Lys	Ala	Leu	Ser	Lys	Ser	Phe	Thr	Met	His	Ser	Leu	Thr	Ile
				290					295					300
Gln	Met	Glu	Ser	Pro	Val	Asp	Gln	Asp	Pro	Asp	Cys	Leu	Phe	Cys
				305					310					315
Glu	Asp	Pro	Cys	Asp										
				320										

<210> 15

<211> 462

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2329339CD1

<400> 15

Met	Ala	Glu	Glu	Gln	Glu	Phe	Thr	Gln	Leu	Cys	Lys	Leu	Pro	Ala
1				5					10					15
Gln	Pro	Ser	His	Pro	His	Cys	Val	Asn	Asn	Thr	Tyr	Arg	Ser	Ala
				20					25					30
Gln	His	Ser	Gln	Ala	Leu	Leu	Arg	Gly	Leu	Leu	Ala	Leu	Arg	Asp
				35					40					45
Ser	Gly	Ile	Leu	Phe	Asp	Val	Val	Leu	Val	Val	Glu	Gly	Arg	His
				50					55					60
Ile	Glu	Ala	His	Arg	Ile	Leu	Leu	Ala	Ala	Ser	Cys	Asp	Tyr	Phe
				65					70					75
Arg	Gly	Met	Phe	Ala	Gly	Gly	Leu	Lys	Glu	Met	Glu	Gln	Glu	Glu
				80					85					90
Val	Leu	Ile	His	Gly	Val	Ser	Tyr	Asn	Ala	Met	Cys	Gln	Ile	Leu
				95					100					105
His	Phe	Ile	Tyr	Thr	Ser	Glu	Leu	Glu	Leu	Ser	Leu	Ser	Asn	Val
				110					115					120
Gln	Glu	Thr	Leu	Val	Ala	Ala	Cys	Gln	Leu	Gln	Ile	Pro	Glu	Ile
				125					130					135
Ile	His	Phe	Cys	Cys	Asp	Phe	Leu	Met	Ser	Trp	Val	Asp	Glu	Glu
				140					145					150
Asn	Ile	Leu	Asp	Val	Tyr	Arg	Leu	Ala	Glu	Leu	Phe	Asp	Leu	Ser
				155					160					165
Arg	Leu	Thr	Glu	Gln	Leu	Asp	Thr	Tyr	Ile	Leu	Lys	Asn	Phe	Val
				170					175					180
Ala	Phe	Ser	Arg	Thr	Asp	Lys	Tyr	Arg	Gln	Leu	Pro	Leu	Glu	Lys
				185					190					195
Val	Tyr	Ser	Leu	Leu	Ser	Ser	Asn	Arg	Leu	Glu	Val	Ser	Cys	Glu
				200					205					210
Thr	Glu	Val	Tyr	Glu	Gly	Ala	Leu	Leu	Tyr	His	Tyr	Ser	Leu	Glu
				215					220					225
Gln	Val	Gln	Ala	Asp	Gln	Ile	Ser	Leu	His	Glu	Pro	Pro	Lys	Leu
				230					235					240
Leu	Glu	Thr	Val	Arg	Phe	Pro	Leu	Met	Glu	Ala	Glu	Val	Leu	Gln
				245					250					255
Arg	Leu	His	Asp	Lys	Leu	Asp	Pro	Ser	Pro	Leu	Arg	Asp	Thr	Val
				260					265					270
Ala	Ser	Gly	Leu	Met	Tyr	His	Arg	Asn	Glu	Ser	Leu	Gln	Pro	Ser
				275					280					285
Leu	Gln	Ser	Pro	Gln	Thr	Glu	Leu	Arg	Ser	Asp	Phe	Gln	Cys	Val
				290					295					300
Val	Gly	Phe	Gly	Gly	Ile	His	Ser	Thr	Pro	Ser	Thr	Val	Leu	Ser
				305					310					315
Asp	Gln	Ala	Lys	Tyr	Leu	Asn	Pro	Leu	Leu	Gly	Glu	Trp	Lys	His
				320					325					330
Phe	Thr	Ala	Ser	Leu	Ala	Pro	Arg	Met	Ser	Asn	Gln	Gly	Ile	Ala
				335					340					345



WO 00/78953

PCT/US00/16668

Leu	Arg	Cys	Ile	Phe	Leu	Val	Asp	Cys	Arg	Tyr	Cys	Gly	Gly	Val
				140					145					150
Arg	Arg	Asn	Leu	Arg	Gln	Ile	Phe	Gln	Ser	Leu	Pro	Pro	Phe	Met
				155					160					165
Asp	Ile	Leu	Leu	Leu	Leu	Phe	Phe	Met	Ile	Ile	Phe	Ala	Ile	
				170				175						180
Leu	Gly	Phe	Tyr	Leu	Phe	Ser	Pro	Asn	Pro	Ser	Asp	Pro	Tyr	Phe
				185					190					195
Ser	Thr	Leu	Glu	Asn	Ser	Ile	Val	Ser	Leu	Phe	Val	Leu	Leu	Thr
				200					205					210
Thr	Ala	Asn	Phe	Pro	Asp	Val	Met	Met	Pro	Ser	Tyr	Ser	Arg	Asn
				215					220					225
Pro	Trp	Ser	Cys	Val	Phe	Phe	Ile	Val	Tyr	Leu	Ser	Ile	Glu	Leu
				230					235					240
Tyr	Phe	Ile	Met	Asn	Leu	Leu	Leu	Ala	Val	Val	Phe	Asp	Thr	Phe
				245					250					255
Asn	Asp	Ile	Glu	Lys	Arg	Lys	Phe	Lys	Ser	Leu	Leu	Leu	His	Lys
				260					265					270
Arg	Thr	Ala	Ile	Gln	His	Ala	Tyr	Arg	Leu	Leu	Ile	Ser	Gln	Arg
				275					280					285
Arg	Pro	Ala	Gly	Ile	Ser	Tyr	Arg	Gln	Phe	Glu	Gly	Leu	Met	Arg
				290					295					300
Phe	Tyr	Lys	Pro	Arg	Met	Ser	Ala	Arg	Glu	Arg	Tyr	Leu	Thr	Phe
				305					310					315
Lys	Ala	Leu	Asn	Gln	Asn	Asn	Thr	Pro	Leu	Leu	Ser	Leu	Lys	Asp
				320					325					330
Phe	Tyr	Asp	Ile	Tyr	Glu	Val	Ala	Ala	Leu	Lys	Trp	Lys	Ala	Lys
				335					340					345
Lys	Asn	Arg	Glu	His	Trp	Phe	Asp	Glu	Leu	Pro	Arg	Thr	Ala	Leu
				350					355					360
Leu	Ile	Phe	Lys	Gly	Ile	Asn	Ile	Leu	Val	Lys	Ser	Lys	Ala	Phe
				365					370					375
Gln	Tyr	Phe	Met	Tyr	Leu	Val	Val	Ala	Val	Asn	Gly	Val	Trp	Ile
				380					385					390
Leu	Val	Glu	Thr	Phe	Met	Leu	Lys	Gly	Gly	Asn	Phe	Phe	Ser	Lys
				395					400					405
His	Val	Pro	Trp	Ser	Tyr	Leu	Val	Phe	Leu	Thr	Ile	Tyr	Gly	Val
				410					415					420
Glu	Leu	Phe	Leu	Lys	Val	Ala	Gly	Leu	Gly	Pro	Val	Glu	Tyr	Leu
				425					430					435
Ser	Ser	Gly	Trp	Asn	Leu	Phe	Asp	Phe	Ser	Val	Thr	Val	Phe	Ala
				440					445					450
Phe	Leu	Gly	Leu	Leu	Ala	Leu	Ala	Leu	Asn	Met	Glu	Pro	Phe	Tyr
				455					460					465
Phe	Ile	Val	Val	Leu	Arg	Pro	Leu	Gln	Leu	Leu	Arg	Leu	Phe	Lys
				470					475					480
Leu	Lys	Glu	Arg	Tyr	Arg	Asn	Val	Leu	Asp	Thr	Met	Phe	Glu	Leu
				485					490					495
Leu	Pro	Arg	Met	Ala	Ser	Leu	Gly	Leu	Thr	Leu	Leu	Ile	Phe	Tyr
				500					505					510
Tyr	Ser	Phe	Ala	Ile	Val	Gly	Met	Glu	Phe	Phe	Cys	Gly	Ile	Val
				515					520					525
Phe	Pro	Asn	Cys	Cys	Asn	Thr	Ser	Thr	Val	Ala	Asp	Ala	Tyr	Arg
				530					535					540
Trp	Arg	Asn	His	Thr	Val	Gly	Asn	Arg	Thr	Val	Val	Glu	Glu	Gly
				545					550					555
Tyr	Tyr	Tyr	Leu	Asn	Asn	Phe	Asp	Asn	Ile	Leu	Asn	Ser	Phe	Val
				560					565					570
Thr	Leu	Phe	Glu	Leu	Thr	Val	Val	Asn	Asn	Trp	Tyr	Ile	Ile	Met
				575					580					585
Glu	Gly	Val	Thr	Ser	Gln	Thr	Ser	His	Trp	Ser	Arg	Leu	Tyr	Phe
				590					595					600
Met	Thr	Phe	Tyr	Ile	Val	Thr	Met	Val	Val	Met	Thr	Ile	Ile	Val
				605					610					615
Ala	Phe	Ile	Leu	Glu	Ala	Phe	Val	Phe	Arg	Met	Asn	Tyr	Ser	Arg
				620					625					630
Lys	Asn	Gln	Asp	Ser	Glu	Val	Asp	Gly	Gly	Ile	Thr	Leu	Glu	Lys
				635					640					645





WO 00/78953

PCT/US00/16668

Lys	Glu	Leu	Leu	Arg	Arg	Leu	Lys	Val	Gln	Asp	Gln	Met	Thr	Lys	
				335					340					345	
Gln	His	Gln	Thr	Arg	Leu	Asp	Ile	Ile	Ser	Glu	Asp	Ile	Ser	Glu	
				350					355					360	
Leu	Gln	Lys	Asn	Gln	Thr	Thr	Ser	Val	Ala	Lys	Ile	Ala	Gln	Tyr	
				365					370					375	
Lys	Arg	Lys	Leu	Met	Asp	Leu	Ser	His	Arg	Thr	Leu	Gln	Val	Leu	
				380					385					390	
Ile	Lys	Gln	Glu	Ile	Gln	Arg	Lys	Ser	Gly	Tyr	Ala	Ile	Gln	Ala	
				395					400					405	
Asp	Glu	Glu	Gln	Leu	Arg	Val	Gln	Leu	Asp	Thr	Ile	Gln	Gly	Glu	
				410					415					420	
Leu	Asn	Ala	Pro	Thr	Gln	Phe	Lys	Gly	Arg	Leu	Asn	Glu	Leu	Met	
				425					430					435	
Ser	Gln	Ile	Arg	Met	Gln	Asn	His	Phe	Gly	Ala	Val	Arg	Ser	Glu	
				440					445					450	
Glu	Arg	Tyr	Tyr	Ile	Asp	Ala	Asp	Leu	Leu	Arg	Glu	Ile	Lys	Gln	
				455					460					465	
His	Leu	Lys	Gln	Gln	Gln	Glu	Gly	Leu	Ser	His	Leu	Ile	Ser	Ile	
				470					475					480	
Ile	Lys	Asp	Asp	Leu	Glu	Asp	Ile	Lys	Leu	Val	Glu	His	Gly	Leu	
				485					490					495	
Asn	Glu	Thr	Ile	His	Ile	Arg	Gly	Gly	Val	Phe	Ser				
				500					505						

<210> 19

<211> 592

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2758310CD1

<400> 19

Met	Trp	Phe	Cys	Gly	Gln	Ser	Thr	Pro	Phe	Gly	Cys	Glu	Leu	His	
1				5					10					15	
Asp	Thr	Cys	Val	Gln	Leu	Cys	His	Phe	His	Ser	Ala	Leu	Leu	His	
				20					25					30	
Arg	Arg	Gln	Lys	Pro	Trp	Pro	Ser	Pro	Ala	Val	Phe	Phe	Arg	Arg	
				35					40					45	
Asn	Val	Arg	Gly	Leu	Pro	Pro	Arg	Phe	Ser	Ser	Pro	Thr	Pro	Leu	
				50					55					60	
Trp	Arg	Lys	Val	Leu	Ser	Thr	Ala	Val	Val	Gly	Ala	Pro	Leu	Leu	
				65					70					75	
Leu	Gly	Ala	Arg	Tyr	Val	Met	Ala	Glu	Ala	Arg	Glu	Lys	Arg	Arg	
				80					85					90	
Met	Arg	Leu	Val	Val	Asp	Gly	Met	Gly	Arg	Phe	Gly	Arg	Ser	Leu	
				95					100					105	
Lys	Val	Gly	Leu	Gln	Ile	Ser	Leu	Asp	Tyr	Trp	Trp	Cys	Thr	Asn	
				110					115					120	
Val	Val	Leu	Arg	Gly	Trp	Lys	Ser	Pro	Gly	Tyr	Leu	Glu	Val	Met	
				125					130					135	
Ser	Ala	Cys	His	Gln	Arg	Ala	Ala	Asp	Ala	Leu	Val	Ala	Gly	Ala	
				140					145					150	
Ile	Ser	Asn	Gly	Gly	Leu	Tyr	Val	Lys	Leu	Gly	Gln	Gly	Leu	Cys	
				155					160					165	
Ser	Phe	Asn	His	Leu	Leu	Pro	Pro	Glu	Tyr	Thr	Arg	Thr	Leu	Arg	
				170					175					180	
Val	Leu	Glu	Asp	Arg	Ala	Leu	Lys	Arg	Gly	Phe	Gln	Glu	Val	Asp	
				185					190					195	
Glu	Leu	Phe	Leu	Glu	Asp	Phe	Gln	Ala	Leu	Pro	His	Glu	Leu	Phe	
				200					205					210	
Gln	Glu	Phe	Asp	Tyr	Gln	Pro	Ile	Ala	Ala	Ala	Ser	Leu	Ala	Gln	
				215					220					225	
Val	His	Arg	Ala	Lys	Leu	His	Asp	Gly	Thr	Ser	Val	Ala	Val	Lys	
				230					235					240	
Val	Gln	Tyr	Ile	Asp	Leu	Arg	Asp	Arg	Phe	Asp	Gly	Asp	Ile	His	
				245					250					255	

WO 00/78953

PCT/US00/16668

Thr	Leu	Glu	Leu	Leu	Leu	Arg	Leu	Val	Glu	Val	Met	His	Pro	Ser	
				260					265						270
Phe	Gly	Phe	Ser	Trp	Val	Leu	Gln	Asp	Leu	Lys	Gly	Thr	Leu	Ala	
				275					280						285
Gln	Glu	Leu	Asp	Phe	Glu	Asn	Glu	Gly	Arg	Asn	Ala	Glu	Arg	Cys	
				290					295						300
Ala	Arg	Glu	Leu	Ala	His	Phe	Pro	Tyr	Val	Val	Val	Pro	Arg	Val	
				305					310						315
His	Trp	Asp	Lys	Ser	Ser	Lys	Arg	Val	Leu	Thr	Ala	Asp	Phe	Cys	
				320					325						330
Ala	Gly	Cys	Lys	Val	Asn	Asp	Val	Glu	Ala	Ile	Arg	Ser	Gln	Gly	
				335					340						345
Leu	Ala	Val	His	Asp	Ile	Ala	Glu	Lys	Leu	Ile	Lys	Ala	Phe	Ala	
				350					355						360
Glu	Gln	Ile	Phe	Tyr	Thr	Gly	Phe	Ile	His	Ser	Asp	Pro	His	Pro	
				365					370						375
Gly	Asn	Val	Leu	Val	Arg	Lys	Gly	Pro	Asp	Gly	Lys	Ala	Glu	Leu	
				380					385						390
Val	Leu	Leu	Asp	His	Gly	Leu	Tyr	Gln	Phe	Leu	Glu	Glu	Lys	Asp	
				395					400						405
Arg	Ala	Ala	Leu	Cys	Gln	Leu	Trp	Arg	Ala	Ile	Ile	Leu	Arg	Asp	
				410					415						420
Asp	Ala	Ala	Met	Arg	Ala	His	Ala	Ala	Ala	Leu	Gly	Val	Gln	Asp	
				425					430						435
Tyr	Leu	Leu	Phe	Ala	Glu	Met	Leu	Met	Gln	Arg	Pro	Val	Arg	Leu	
				440					445						450
Gly	Gln	Leu	Trp	Gly	Ser	His	Leu	Leu	Ser	Arg	Glu	Glu	Ala	Ala	
				455					460						465
Tyr	Met	Val	Asp	Met	Ala	Arg	Glu	Arg	Phe	Glu	Ala	Val	Met	Ala	
				470					475						480
Val	Leu	Arg	Glu	Leu	Pro	Arg	Pro	Met	Leu	Leu	Val	Leu	Arg	Asn	
				485					490						495
Ile	Asn	Thr	Val	Arg	Ala	Ile	Asn	Val	Ala	Leu	Gly	Ala	Pro	Val	
				500					505						510
Asp	Arg	Tyr	Phe	Leu	Met	Ala	Lys	Arg	Ala	Val	Arg	Gly	Trp	Ser	
				515					520						525
Arg	Leu	Ala	Gly	Ala	Thr	Tyr	Arg	Gly	Val	Tyr	Gly	Thr	Ser	Leu	
				530					535						540
Leu	Arg	His	Ala	Lys	Val	Val	Trp	Glu	Met	Leu	Lys	Phe	Glu	Val	
				545					550						555
Ala	Leu	Arg	Leu	Glu	Thr	Leu	Ala	Met	Arg	Leu	Thr	Ala	Leu	Leu	
				560					565						570
Ala	Arg	Ala	Leu	Val	His	Leu	Ser	Leu	Val	Pro	Pro	Ala	Glu	Glu	
				575					580						585
Leu	Tyr	Gln	Tyr	Leu	Glu	Thr									
				590											

<210> 20

<211> 841

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 2762348CD1

<400> 20

Met	Ala	Ser	Val	Phe	Arg	Ser	Glu	Glu	Met	Cys	Leu	Ser	Gln	Leu	
1				5					10					15	
Phe	Leu	Gln	Val	Glu	Ala	Ala	Tyr	Cys	Cys	Val	Ala	Glu	Leu	Gly	
				20					25					30	
Glu	Leu	Gly	Leu	Val	Gln	Phe	Lys	Asp	Leu	Asn	Met	Asn	Val	Asn	
				35					40					45	
Ser	Phe	Gln	Arg	Lys	Phe	Val	Asn	Glu	Val	Arg	Arg	Cys	Glu	Ser	
				50					55					60	
Leu	Glu	Arg	Ile	Leu	Arg	Phe	Leu	Glu	Asp	Glu	Met	Gln	Asn	Glu	
				65					70					75	
Ile	Val	Val	Gln	Leu	Leu	Glu	Lys	Ser	Pro	Leu	Thr	Pro	Leu	Pro	
				80					85					90	

WO 00/78953

PCT/US00/16668

Arg	Glu	Met	Ile	Thr	Leu	Glu	Thr	Val	Leu	Glu	Lys	Leu	Glu	Gly
				95					100					105
Glu	Leu	Gln	Glu	Ala	Asn	Gln	Asn	Gln	Gln	Ala	Leu	Lys	Gln	Ser
				110					115					120
Phe	Leu	Glu	Leu	Thr	Glu	Leu	Lys	Tyr	Leu	Leu	Lys	Lys	Thr	Gln
				125					130					135
Asp	Phe	Phe	Glu	Thr	Glu	Thr	Asn	Leu	Ala	Asp	Asp	Phe	Phe	Thr
				140					145					150
Glu	Asp	Thr	Ser	Gly	Leu	Leu	Glu	Leu	Lys	Ala	Val	Pro	Ala	Tyr
				155					160					165
Met	Thr	Gly	Lys	Leu	Gly	Phe	Ile	Ala	Gly	Cys	Asp	Pro	Thr	Gly
				170					175					180
Lys	Arg	Met	Ala	Ser	Phe	Glu	Arg	Leu	Leu	Trp	Arg	Val	Cys	Arg
				185					190					195
Gly	Asn	Val	Tyr	Leu	Lys	Phe	Ser	Glu	Met	Asp	Ala	Pro	Leu	Glu
				200					205					210
Asp	Pro	Val	Thr	Lys	Glu	Glu	Ile	Gln	Lys	His	Ile	Phe	Ile	Ile
				215					220					225
Phe	Tyr	Gln	Gly	Glu	Gln	Leu	Arg	Gln	Lys	Ile	Lys	Lys	Ile	Cys
				230					235					240
Asp	Gly	Phe	Arg	Ala	Thr	Val	Tyr	Pro	Cys	Pro	Glu	Pro	Ala	Val
				245					250					255
Glu	Arg	Arg	Glu	Met	Leu	Glu	Ser	Val	Asn	Val	Arg	Leu	Glu	Asp
				260					265					270
Leu	Ile	Thr	Val	Ile	Thr	Gln	Thr	Glu	Ser	His	Arg	Gln	Arg	Leu
				275					280					285
Leu	Gln	Glu	Ala	Ala	Ala	Asn	Trp	His	Ser	Trp	Leu	Ile	Lys	Val
				290					295					300
Gln	Lys	Met	Lys	Ala	Val	Tyr	His	Ile	Leu	Asn	Met	Cys	Asn	Ile
				305					310					315
Asp	Val	Thr	Gln	Gln	Cys	Val	Ile	Ala	Glu	Ile	Trp	Phe	Pro	Val
				320					325					330
Ala	Asp	Ala	Thr	Arg	Ile	Lys	Arg	Ala	Leu	Glu	Gln	Gly	Met	Glu
				335					340					345
Leu	Ser	Gly	Ser	Ser	Met	Ala	Pro	Ile	Met	Thr	Thr	Val	Gln	Ser
				350					355					360
Lys	Thr	Ala	Pro	Pro	Thr	Phe	Asn	Arg	Thr	Asn	Lys	Phe	Thr	Ala
				365					370					375
Gly	Phe	Gln	Asn	Ile	Val	Asp	Ala	Tyr	Gly	Val	Gly	Ser	Tyr	Arg
				380					385					390
Glu	Ile	Asn	Pro	Ala	Pro	Tyr	Thr	Ile	Ile	Thr	Phe	Pro	Phe	Leu
				395					400					405
Phe	Ala	Val	Met	Phe	Gly	Asp	Cys	Gly	His	Gly	Thr	Val	Met	Leu
				410					415					420
Leu	Ala	Ala	Leu	Trp	Met	Ile	Leu	Asn	Glu	Arg	Arg	Leu	Leu	Ser
				425					430					435
Gln	Lys	Thr	Asp	Asn	Glu	Ile	Trp	Asn	Thr	Phe	Phe	His	Gly	Arg
				440					445					450
Tyr	Leu	Ile	Leu	Leu	Met	Gly	Ile	Phe	Ser	Ile	Tyr	Thr	Gly	Leu
				455					460					465
Ile	Tyr	Asn	Asp	Cys	Phe	Ser	Lys	Ser	Leu	Asn	Ile	Phe	Gly	Ser
				470					475					480
Ser	Trp	Ser	Val	Gln	Pro	Met	Phe	Arg	Asn	Gly	Thr	Trp	Asn	Thr
				485					490					495
His	Val	Met	Glu	Glu	Ser	Leu	Tyr	Leu	Gln	Leu	Asp	Pro	Ala	Ile
				500					505					510
Pro	Gly	Val	Tyr	Phe	Gly	Asn	Pro	Tyr	Pro	Phe	Gly	Ile	Asp	Pro
				515					520					525
Ile	Trp	Asn	Leu	Ala	Ser	Asn	Lys	Leu	Thr	Phe	Leu	Asn	Ser	Tyr
				530					535					540
Lys	Met	Lys	Met	Ser	Val	Ile	Leu	Gly	Ile	Val	Gln	Met	Val	Phe
				545					550					555
Gly	Val	Ile	Leu	Ser	Leu	Phe	Asn	His	Ile	Tyr	Phe	Arg	Arg	Thr
				560					565					570
Leu	Asn	Ile	Ile	Leu	Gln	Phe	Ile	Pro	Glu	Met	Ile	Phe	Ile	Leu
				575					580					585
Cys	Leu	Phe	Gly	Tyr	Leu	Val	Phe	Met	Ile	Ile	Phe	Lys	Trp	Cys
				590					595					600

WO 00/78953

PCT/US00/16668

Cys	Phe	Asp	Val	His	Val	Ser	Gln	His	Ala	Pro	Ser	Ile	Leu	Ile
				605					610					615
His	Phe	Ile	Asn	Met	Phe	Leu	Phe	Asn	Tyr	Ser	Asp	Ser	Ser	Asn
				620					625					630
Ala	Pro	Leu	Tyr	Lys	His	Gln	Gln	Glu	Val	Gln	Ser	Phe	Phe	Val
				635					640					645
Val	Met	Ala	Leu	Ile	Ser	Val	Pro	Trp	Met	Leu	Leu	Ile	Lys	Pro
				650					655					660
Phe	Ile	Leu	Arg	Ala	Ser	His	Arg	Lys	Ser	Gln	Leu	Gln	Ala	Ser
				665					670					675
Arg	Ile	Gln	Glu	Asp	Ala	Thr	Glu	Asn	Ile	Glu	Gly	Asp	Ser	Ser
				680					685					690
Ser	Pro	Ser	Ser	Arg	Ser	Gly	Gln	Arg	Thr	Ser	Ala	Asp	Thr	His
				695					700					705
Gly	Ala	Leu	Asp	Asp	His	Gly	Glu	Glu	Phe	Asn	Phe	Gly	Asp	Val
				710					715					720
Phe	Val	His	Gln	Ala	Ile	His	Thr	Ile	Glu	Tyr	Cys	Leu	Gly	Cys
				725					730					735
Ile	Ser	Asn	Thr	Ala	Ser	Tyr	Leu	Arg	Leu	Trp	Ala	Leu	Ser	Leu
				740					745					750
Ala	His	Ala	Gln	Leu	Ser	Glu	Val	Leu	Trp	Thr	Met	Val	Met	Asn
				755					760					765
Ser	Gly	Leu	Gln	Thr	Arg	Gly	Trp	Gly	Gly	Ile	Val	Gly	Val	Phe
				770					775					780
Ile	Ile	Phe	Ala	Val	Phe	Ala	Val	Leu	Thr	Val	Ala	Ile	Leu	Leu
				785					790					795
Ile	Met	Glu	Gly	Leu	Ser	Ala	Phe	Leu	His	Ala	Leu	Arg	Leu	His
				800					805					810
Trp	Val	Glu	Phe	Gln	Asn	Lys	Phe	Tyr	Val	Gly	Asp	Gly	Tyr	Lys
				815					820					825
Phe	Ser	Pro	Phe	Ser	Phe	Lys	His	Ile	Leu	Asp	Gly	Thr	Ala	Glu
				830					835					840

Glu

<210> 21

<211> 253

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 3715961CD1

<400> 21

Met	Ser	Glu	Cys	Pro	Leu	Ile	Leu	Tyr	Ile	His	Lys	His	Ile	Asp
1				5					10					15
Thr	Tyr	Ser	Gln	Ser	Tyr	Leu	Phe	Asn	Asp	Leu	Phe	Tyr	Pro	Val
				20					25					30
Tyr	Ser	Gly	Gly	Arg	Met	Val	Thr	Tyr	Glu	His	Leu	Arg	Glu	Val
				35					40					45
Val	Phe	Gly	Lys	Ser	Glu	Asp	Glu	His	Tyr	Pro	Leu	Trp	Lys	Ser
				50					55					60
Val	Ile	Gly	Gly	Met	Met	Ala	Gly	Val	Ile	Gly	Gln	Phe	Leu	Ala
				65					70					75
Asn	Pro	Thr	Asp	Leu	Val	Lys	Val	Gln	Met	Gln	Met	Glu	Gly	Lys
				80					85					90
Arg	Lys	Leu	Glu	Gly	Lys	Pro	Leu	Arg	Phe	Arg	Gly	Val	His	His
				95					100					105
Ala	Phe	Ala	Lys	Ile	Leu	Ala	Glu	Gly	Gly	Ile	Arg	Gly	Leu	Trp
				110					115					120
Ala	Gly	Trp	Val	Pro	Asn	Ile	Gln	Arg	Ala	Ala	Leu	Val	Asn	Met
				125					130					135
Gly	Asp	Leu	Thr	Thr	Tyr	Asp	Thr	Val	Lys	His	Tyr	Leu	Val	Leu
				140					145					150
Asn	Thr	Pro	Leu	Glu	Asp	Asn	Ile	Met	Thr	His	Gly	Leu	Ser	Ser
				155					160					165
Leu	Cys	Ser	Gly	Leu	Val	Ala	Ser	Ile	Leu	Gly	Thr	Pro	Ala	Asp
				170					175					180

WO 00/78953

PCT/US00/16668

Val	Ile	Lys	Ser	Arg	Ile	Met	Asn	Gln	Pro	Arg	Asp	Lys	Gln	Gly
				185					190					195
Arg	Gly	Leu	Leu	Tyr	Lys	Ser	Ser	Thr	Asp	Cys	Leu	Ile	Gln	Ala
				200					205					210
Val	Gln	Gly	Glu	Gly	Phe	Met	Ser	Leu	Tyr	Lys	Gly	Phe	Leu	Pro
				215					220					225
Ser	Trp	Leu	Arg	Met	Thr	Pro	Trp	Ser	Met	Val	Phe	Trp	Leu	Thr
				230					235					240
Tyr	Glu	Lys	Ile	Arg	Glu	Met	Ser	Gly	Val	Ser	Pro	Phe		
				245					250					

<210> 22  
 <211> 229  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5108194CD1

Met	Gly	Asn	Gly	Val	Lys	Glu	Gly	Pro	Val	Arg	Leu	His	Glu	Asp
1				5					10					15
Ala	Glu	Ala	Val	Leu	Ser	Ser	Ser	Val	Ser	Ser	Lys	Arg	Asp	His
				20					25					30
Arg	Gln	Val	Leu	Ser	Ser	Leu	Leu	Ser	Gly	Ala	Leu	Ala	Gly	Ala
				35					40					45
Leu	Ala	Lys	Thr	Ala	Val	Ala	Pro	Leu	Asp	Arg	Thr	Lys	Ile	Ile
				50					55					60
Phe	Gln	Val	Ser	Ser	Lys	Arg	Phe	Ser	Ala	Lys	Glu	Ala	Phe	Arg
				65					70					75
Val	Leu	Tyr	Tyr	Thr	Tyr	Leu	Asn	Glu	Gly	Phe	Leu	Ser	Leu	Trp
				80					85					90
Arg	Gly	Asn	Ser	Ala	Thr	Met	Val	Arg	Val	Val	Pro	Tyr	Ala	Ala
				95					100					105
Ile	Gln	Phe	Ser	Ala	His	Glu	Glu	Tyr	Lys	Arg	Ile	Leu	Gly	Ser
				110					115					120
Tyr	Tyr	Gly	Phe	Arg	Gly	Glu	Ala	Leu	Pro	Pro	Trp	Pro	Arg	Leu
				125					130					135
Phe	Ala	Gly	Ala	Leu	Ala	Gly	Thr	Thr	Ala	Ala	Ser	Leu	Thr	Tyr
				140					145					150
Pro	Leu	Asp	Leu	Val	Arg	Ala	Arg	Met	Ala	Val	Thr	Pro	Lys	Glu
				155					160					165
Met	Tyr	Ser	Asn	Ile	Phe	His	Val	Phe	Ile	Arg	Ile	Ser	Arg	Glu
				170					175					180
Glu	Gly	Leu	Lys	Thr	Leu	Tyr	His	Gly	Phe	Met	Pro	Thr	Val	Leu
				185					190					195
Gly	Val	Ile	Pro	Tyr	Ala	Gly	Leu	Ser	Phe	Phe	Thr	Tyr	Glu	Thr
				200					205					210
Leu	Lys	Ser	Leu	His	Arg	Glu	Tyr	Ser	Gly	Arg	Lys	Leu	Ile	Pro
				215					220					225
Phe	Ser	Glu	Gly											

<210> 23  
 <211> 170  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 5503122CD1

Met	Tyr	Asp	Asn	Leu	Tyr	Leu	His	Gly	Ile	Glu	Asp	Ser	Glu	Ala
1				5					10					15
Gly	Ser	Ala	Asp	Ser	Tyr	Thr	Ser	Arg	Pro	Ser	Asp	Ser	Asp	Val
				20					25					30
Ser	Leu	Glu	Glu	Asp	Arg	Glu	Ala	Ile	Arg	Gln	Glu	Arg	Glu	Gln
				35					40					45



WO 00/78953

PCT/US00/16668

Asp	Ser	Thr	Ala	Val	Ala	Leu	Asn	Arg	Glu	Glu	Asp	Phe	Lys	Ala
				305					310					315
Thr	Glu	Ile	Ile	Glu	Pro	Ser	Lys	Gln	Asp	Lys	Pro	Leu	Ile	Glu
				320					325					330
Lys	Leu	Ala	Glu	Ile	Tyr	Val	Asn	Ser	Ser	Phe	Tyr	Lys	Glu	Thr
				335					340					345
Lys	Ala	Glu	Leu	His	Gln	Leu	Ser	Gly	Gly	Glu	Lys	Lys	Lys	Lys
				350					355					360
Ile	Thr	Val	Phe	Lys	Glu	Ile	Ser	Tyr	Thr	Thr	Ser	Phe	Cys	His
				365					370					375
Gln	Leu	Arg	Trp	Val	Ser	Lys	Arg	Ser	Phe	Lys	Asn	Leu	Leu	Gly
				380					385					390
Asn	Pro	Gln	Ala	Ser	Ile	Ala	Gln	Ile	Ile	Val	Thr	Val	Val	Leu
				395					400					405
Gly	Leu	Val	Ile	Gly	Ala	Ile	Tyr	Phe	Gly	Leu	Lys	Asn	Asp	Ser
				410					415					420
Thr	Gly	Ile	Gln	Asn	Arg	Ala	Gly	Val	Leu	Phe	Phe	Leu	Thr	Thr
				425					430					435
Asn	Gln	Cys	Phe	Ser	Ser	Val	Ser	Ala	Val	Glu	Leu	Phe	Val	Val
				440					445					450
Glu	Lys	Lys	Leu	Phe	Ile	His	Glu	Tyr	Ile	Ser	Gly	Tyr	Tyr	Arg
				455					460					465
Val	Ser	Ser	Tyr	Phe	Leu	Gly	Lys	Leu	Leu	Ser	Asp	Leu	Leu	Pro
				470					475					480
Met	Arg	Met	Leu	Pro	Ser	Ile	Ile	Phe	Thr	Cys	Ile	Val	Tyr	Phe
				485					490					495
Met	Leu	Gly	Leu	Lys	Pro	Lys	Ala	Asp	Ala	Phe	Phe	Val	Met	Met
				500					505					510
Phe	Thr	Leu	Met	Met	Val	Ala	Tyr	Ser	Ala	Ser	Ser	Met	Ala	Leu
				515					520					525
Ala	Ile	Ala	Ala	Gly	Gln	Ser	Val	Val	Ser	Val	Ala	Thr	Leu	Leu
				530					535					540
Met	Thr	Ile	Cys	Phe	Val	Phe	Met	Met	Ile	Phe	Ser	Gly	Leu	Leu
				545					550					555
Val	Asn	Leu	Thr	Thr	Ile	Ala	Ser	Trp	Leu	Ser	Trp	Leu	Gln	Tyr
				560					565					570
Phe	Ser	Ile	Pro	Arg	Tyr	Gly	Phe	Thr	Ala	Leu	Gln	His	Asn	Glu
				575					580					585
Phe	Leu	Gly	Gln	Asn	Phe	Cys	Pro	Gly	Leu	Asn	Ala	Thr	Gly	Asn
				590					595					600
Asn	Pro	Cys	Asn	Tyr	Ala	Thr	Cys	Thr	Gly	Glu	Glu	Tyr	Leu	Val
				605					610					615
Lys	Gln	Gly	Ile	Asp	Leu	Ser	Pro	Trp	Gly	Leu	Trp	Lys	Asn	His
				620					625					630
Val	Ala	Leu	Ala	Cys	Met	Ile	Val	Ile	Phe	Leu	Thr	Ile	Ala	Tyr
				635					640					645
Leu	Lys	Leu	Leu	Phe	Leu	Lys	Lys	Tyr	Ser					
				650					655					

<210> 25

<211> 184

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 5593114CD1

<400> 25

Met	Trp	Val	Phe	Gly	Tyr	Gly	Ser	Leu	Ile	Trp	Lys	Val	Asp	Phe
1				5					10					15
Pro	Tyr	Gln	Asp	Lys	Leu	Val	Gly	Tyr	Ile	Thr	Asn	Tyr	Ser	Arg
				20					25					30
Arg	Phe	Trp	Gln	Gly	Ser	Thr	Asp	His	Arg	Gly	Val	Pro	Gly	Lys
				35					40					45
Pro	Gly	Arg	Val	Val	Thr	Leu	Val	Glu	Asp	Pro	Ala	Gly	Cys	Val
				50					55					60
Trp	Gly	Val	Ala	Tyr	Arg	Leu	Pro	Val	Gly	Lys	Glu	Glu	Glu	Val
				65					70					75

WO 00/78953

PCT/US00/16668

Lys Ala Tyr Leu Asp Phe Arg Glu Lys Gly Gly Tyr Arg Thr Thr  
 80 85 90  
 Thr Val Ile Phe Tyr Pro Lys Asp Pro Thr Thr Lys Pro Phe Ser  
 95 100 105  
 Val Leu Leu Tyr Ile Gly Thr Cys Asp Asn Pro Asp Tyr Leu Gly  
 110 115 120  
 Pro Ala Pro Leu Glu Asp Ile Ala Glu Gln Ile Phe Asn Ala Ala  
 125 130 135  
 Gly Pro Ser Gly Arg Asn Thr Glu Tyr Leu Phe Glu Leu Ala Asn  
 140 145 150  
 Ser Ile Arg Asn Leu Val Pro Glu Glu Ala Asp Glu His Leu Phe  
 155 160 165  
 Ala Leu Glu Lys Leu Val Lys Glu Arg Leu Glu Gly Lys Gln Asn  
 170 175 180  
 Leu Asn Cys Ile

&lt;210&gt; 26

&lt;211&gt; 154

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 044775CD1

&lt;400&gt; 26

Met Gly Ala Phe Glu Cys Val Arg Lys Val Tyr Gln Thr Asp Gly  
 1 5 10 15  
 Leu Lys Gly Phe Tyr Arg Gly Met Ser Ala Ser Tyr Ala Gly Ile  
 20 25 30  
 Ser Glu Thr Val Ile His Phe Val Ile Tyr Glu Ser Ile Lys Gln  
 35 40 45  
 Lys Leu Leu Glu Tyr Lys Thr Ala Ser Thr Met Glu Asn Asp Glu  
 50 55 60  
 Glu Ser Val Lys Glu Ala Ser Asp Phe Val Gly Met Met Leu Ala  
 65 70 75  
 Ala Ala Thr Ser Lys Thr Cys Ala Thr Thr Ile Ala Tyr Pro His  
 80 85 90  
 Glu Val Val Arg Thr Arg Leu Arg Glu Glu Gly Thr Lys Tyr Arg  
 95 100 105  
 Ser Phe Phe Gln Thr Leu Ser Leu Leu Val Gln Glu Glu Gly Tyr  
 110 115 120  
 Gly Ser Leu Tyr Arg Gly Leu Thr Thr His Leu Val Arg Gln Ile  
 125 130 135  
 Pro Asn Thr Ala Ile Met Met Ala Thr Tyr Glu Leu Val Val Tyr  
 140 145 150  
 Leu Leu Asn Gly

&lt;210&gt; 27

&lt;211&gt; 438

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 116588CD1

&lt;400&gt; 27

Met Leu Leu Val Thr Pro Arg Pro Glu Arg Gly Gly Arg Gly Thr  
 1 5 10 15  
 Glu Leu Gly Glu Phe Cys Gly Thr Pro Leu Leu Phe Ser Ser Tyr  
 20 25 30  
 Phe Cys Tyr Asp Asn Pro Ala Ala Leu Gln Thr Gln Val Lys Arg  
 35 40 45  
 Asp Met Gln Val Asn Thr Thr Lys Phe Met Leu Leu Tyr Ala Trp  
 50 55 60  
 Tyr Ser Trp Pro Asn Val Val Leu Cys Phe Phe Gly Gly Phe Leu  
 65 70 75



WO 00/78953

PCT/US00/16668

```

Ile Asp Arg Val Phe Gly Ile Arg Trp Gly Thr Ile Ile Phe Ser
80 85 90
Cys Phe Val Cys Ile Gly Gln Val Val Phe Ala Leu Gly Gly Ile
95 100 105
Phe Asn Ala Phe Trp Leu Met Glu Phe Gly Arg Phe Val Phe Gly
110 115 120
Ile Gly Gly Glu Ser Leu Ala Val Ala Gln Asn Thr Tyr Ala Val
125 130 135
Ser Trp Phe Lys Gly Lys Glu Leu Asn Leu Val Phe Gly Leu Gln
140 145 150
Leu Ser Met Ala Arg Ile Gly Ser Thr Val Asn Met Asn Leu Met
155 160 165
Gly Trp Leu Tyr Ser Lys Ile Glu Ala Leu Leu Gly Ser Ala Gly
170 175 180
His Thr Thr Leu Gly Ile Thr Leu Met Ile Gly Gly Val Thr Cys
185 190 195
Ile Leu Ser Leu Ile Cys Ala Leu Ala Leu Ala Tyr Leu Asp Gln
200 205 210
Arg Ala Glu Arg Ile Leu His Lys Glu Gln Gly Lys Thr Gly Glu
215 220 225
Val Ile Lys Leu Thr Asp Val Lys Asp Phe Ser Leu Pro Leu Trp
230 235 240
Leu Ile Phe Ile Ile Cys Val Cys Tyr Tyr Val Ala Val Phe Pro
245 250 255
Phe Ile Gly Leu Gly Lys Val Phe Phe Thr Glu Lys Phe Gly Phe
260 265 270
Ser Ser Gln Ala Ala Ser Ala Ile Asn Ser Val Val Tyr Val Ile
275 280 285
Ser Ala Pro Met Ser Pro Val Phe Gly Leu Leu Val Asp Lys Thr
290 295 300
Gly Lys Asn Ile Ile Trp Val Leu Cys Ala Val Ala Ala Thr Leu
305 310 315
Val Ser His Met Met Leu Ala Phe Thr Met Trp Asn Pro Trp Ile
320 325 330
Ala Met Cys Leu Leu Gly Leu Ser Tyr Ser Leu Leu Ala Cys Ala
335 340 345
Leu Trp Pro Met Val Ala Phe Val Val Pro Glu His Gln Leu Gly
350 355 360
Thr Ala Tyr Gly Phe Met Gln Ser Ile Gln Asn Leu Gly Leu Ala
365 370 375
Ile Ile Ser Ile Ile Ala Gly Met Ile Leu Asp Ser Arg Gly Tyr
380 385 390
Leu Phe Leu Glu Val Phe Phe Ile Ala Cys Val Ser Leu Ser Leu
395 400 405
Leu Ser Val Val Leu Leu Tyr Leu Val Asn Arg Ala Gln Gly Gly
410 415 420
Asn Leu Asn Tyr Ser Ala Arg Gln Arg Glu Glu Ile Lys Phe Ser
425 430 435
His Thr Glu

```

<210> 28

<211> 237

<212> PRT

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 875369CD1

<400> 28

```

Met Ala His Val Gly Ser Arg Lys Arg Ser Arg Ser Arg Ser Arg
1 5 10 15
Ser Arg Gly Arg Gly Ser Glu Lys Arg Lys Lys Lys Ser Arg Lys
20 25 30
Asp Thr Ser Arg Asn Cys Ser Ala Ser Thr Ser Gln Gly Arg Lys
35 40 45
Ala Ser Thr Ala Pro Gly Ala Glu Ala Ser Pro Ser Pro Cys Ile
50 55 60

```







WO 00/78953

PCT/US00/16668

				185					190					195
Phe	Gly	Val	Ser	Gly	Phe	Cys	Leu	Leu	Phe	Thr	Phe	Leu	Glu	Leu
				200					205					210
Val	Leu	Leu	Gly	Leu	Gly	Arg	Trp	Trp	Arg	Thr	Trp	Lys	His	Lys
				215					220					225
Ser	Ser	Ser	Ser	Lys	Tyr	Phe	Leu	Thr	Ser	Glu	Ser	Thr	Arg	Arg
				230					235					240
His	Lys	Lys	Ala	Thr	Asp	Ser	Leu	Pro	Val	Val	Glu	Thr	Lys	Glu
				245					250					255
Gln	Phe	Gln	Glu	Ala	Val	Pro	Gly	Arg	Ser	Leu	Ala	Gln	Glu	Lys
				260					265					270
Gln	Arg	Pro	Val	Gly	Pro	Arg	Asp	Ala						
				275										

<210> 32  
 <211> 154  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2219267CD1

<400>	32													
Met	Val	Thr	Gly	Leu	Ala	Ser	Leu	Leu	Leu	Leu	Ala	Gly	Ala	Gln
1				5					10					15
Tyr	Leu	Pro	Gly	Trp	Thr	Val	Leu	Phe	Leu	Ser	Val	Leu	Gly	Leu
				20					25					30
Leu	Ala	Ser	Arg	Ala	Val	Ser	Ala	Leu	Ser	Ser	Leu	Phe	Ala	Ala
				35					40					45
Glu	Val	Phe	Pro	Thr	Val	Ile	Arg	Gly	Ala	Gly	Leu	Gly	Leu	Val
				50					55					60
Leu	Gly	Ala	Gly	Phe	Leu	Gly	Gln	Ala	Ala	Gly	Pro	Leu	Asp	Thr
				65					70					75
Leu	His	Gly	Arg	Gln	Gly	Phe	Phe	Leu	Gln	Gln	Val	Val	Phe	Ala
				80					85					90
Ser	Leu	Ala	Val	Leu	Ala	Leu	Leu	Cys	Val	Leu	Leu	Leu	Pro	Glu
				95					100					105
Ser	Arg	Ser	Arg	Gly	Leu	Pro	Gln	Ser	Leu	Gln	Asp	Ala	Asp	Arg
				110					115					120
Leu	Arg	Arg	Ser	Pro	Leu	Leu	Arg	Gly	Arg	Pro	Arg	Gln	Asp	His
				125					130					135
Leu	Pro	Leu	Leu	Pro	Pro	Ser	Asn	Ser	Tyr	Trp	Ala	Gly	His	Thr
				140					145					150
Pro	Glu	Gln	His											

<210> 33  
 <211> 289  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2308629CD1

<400>	33													
Met	Val	Ala	Gly	Ala	Val	Ala	Gly	Ile	Leu	Glu	His	Cys	Val	Met
1				5					10					15
Tyr	Pro	Ile	Asp	Cys	Val	Lys	Thr	Arg	Met	Gln	Ser	Leu	Gln	Pro
				20					25					30
Asp	Pro	Ala	Ala	Arg	Tyr	Arg	Asn	Val	Leu	Glu	Ala	Leu	Trp	Arg
				35					40					45
Ile	Ile	Arg	Thr	Glu	Gly	Leu	Trp	Arg	Pro	Met	Arg	Gly	Leu	Asn
				50					55					60
Val	Thr	Ala	Thr	Gly	Ala	Gly	Pro	Ala	His	Ala	Leu	Tyr	Phe	Ala
				65					70					75
Cys	Tyr	Glu	Lys	Leu	Lys	Lys	Thr	Leu	Ser	Asp	Val	Ile	His	Pro
				80					85					90
Gly	Gly	Asn	Ser	His	Ile	Ala	Asn	Gly	Ala	Ala	Gly	Cys	Val	Ala

WO 00/78953

PCT/US00/16668

				95					100					105
Thr	Leu	Leu	His	Asp	Ala	Ala	Met	Asn	Pro	Ala	Glu	Val	Val	Lys
				110					115					120
Gln	Arg	Met	Gln	Met	Tyr	Asn	Ser	Pro	Tyr	His	Arg	Val	Thr	Asp
				125					130					135
Cys	Val	Arg	Ala	Val	Trp	Gln	Asn	Glu	Gly	Ala	Gly	Ala	Phe	Tyr
				140					145					150
Arg	Ser	Tyr	Thr	Thr	Gln	Leu	Thr	Met	Asn	Val	Pro	Phe	Gln	Ala
				155					160					165
Ile	His	Phe	Met	Thr	Tyr	Glu	Phe	Leu	Gln	Glu	His	Phe	Asn	Pro
				170					175					180
Gln	Arg	Arg	Tyr	Asn	Pro	Ser	Ser	His	Val	Leu	Ser	Gly	Ala	Cys
				185					190					195
Ala	Gly	Ala	Val	Ala	Ala	Ala	Ala	Thr	Thr	Pro	Leu	Asp	Val	Cys
				200					205					210
Lys	Thr	Leu	Leu	Asn	Thr	Gln	Glu	Ser	Leu	Ala	Leu	Asn	Ser	His
				215					220					225
Ile	Thr	Gly	His	Ile	Thr	Gly	Met	Ala	Ser	Ala	Phe	Arg	Thr	Val
				230					235					240
Tyr	Gln	Val	Gly	Gly	Val	Thr	Ala	Tyr	Phe	Arg	Gly	Val	Gln	Ala
				245					250					255
Arg	Val	Ile	Tyr	Gln	Ile	Pro	Ser	Thr	Ala	Ile	Ala	Trp	Ser	Val
				260					265					270
Tyr	Glu	Phe	Phe	Lys	Tyr	Leu	Ile	Thr	Lys	Arg	Gln	Glu	Glu	Trp
				275					280					285

<210> 34  
 <211> 300  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2660038CD1

<400> 34

Met	Asp	Phe	Leu	Met	Ser	Gly	Leu	Ala	Ala	Cys	Gly	Ala	Cys	Val
1				5					10					15
Phe	Thr	Asn	Pro	Leu	Glu	Val	Val	Lys	Thr	Arg	Met	Gln	Leu	Gln
				20					25					30
Gly	Glu	Leu	Gln	Ala	Pro	Gly	Thr	Tyr	Gln	Arg	His	Tyr	Arg	Asn
				35					40					45
Val	Phe	His	Ala	Phe	Ile	Thr	Ile	Gly	Lys	Val	Asp	Gly	Leu	Ala
				50					55					60
Ala	Leu	Gln	Lys	Gly	Leu	Ala	Pro	Ala	Leu	Leu	Tyr	Gln	Phe	Leu
				65					70					75
Met	Asn	Gly	Ile	Arg	Leu	Gly	Thr	Tyr	Gly	Leu	Ala	Glu	Ala	Gly
				80					85					90
Gly	Tyr	Leu	His	Thr	Ala	Glu	Ala	Thr	His	Ser	Pro	Ala	Arg	Ser
				95					100					105
Ala	Ala	Ala	Gly	Ala	Met	Ala	Gly	Val	Met	Gly	Ala	Tyr	Leu	Gly
				110					115					120
Ser	Pro	Ile	Tyr	Met	Val	Lys	Thr	His	Leu	Gln	Ala	Gln	Ala	Ala
				125					130					135
Ser	Glu	Ile	Ala	Val	Gly	His	Gln	Tyr	Lys	His	Gln	Gly	Met	Phe
				140					145					150
Gln	Ala	Leu	Thr	Glu	Ile	Gly	Gln	Lys	His	Gly	Leu	Val	Gly	Leu
				155					160					165
Trp	Arg	Gly	Ala	Leu	Gly	Gly	Leu	Pro	Arg	Val	Ile	Val	Gly	Ser
				170					175					180
Ser	Thr	Gln	Leu	Cys	Thr	Phe	Ser	Ser	Thr	Lys	Asp	Leu	Leu	Ser
				185					190					195
Gln	Trp	Glu	Ile	Phe	Pro	Pro	Gln	Ser	Trp	Lys	Leu	Ala	Leu	Val
				200					205					210
Ala	Ala	Met	Met	Ser	Gly	Ile	Ala	Val	Val	Leu	Ala	Met	Ala	Pro
				215					220					225
Phe	Asp	Val	Ala	Cys	Thr	Arg	Leu	Tyr	Asn	Gln	Pro	Thr	Asp	Ala

WO 00/78953

PCT/US00/16668

Gln Gly Lys Gly	230	Leu Met Tyr Arg Gly	235	Ile Leu Asp Ala Leu	240
	245		250		255
Gln Thr Ala Arg	260	Glu Gly Ile Phe	265	Met Tyr Lys Gly	270
	275		280		285
Gly Ala Ser Tyr	290	Arg Leu Gly Pro	295	Thr Ile Leu Ser	300
Phe Phe Trp Asp		Leu Arg Ser Leu		Tyr Thr Asp Thr	

<210> 35  
 <211> 382  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2670745CD1

<400> 35

Met Leu Arg Trp Thr	Val His Leu Glu Gly	Gly Pro Arg Arg	Val
1	5	10	15
Asn His Ala Ala Val	Ala Val Gly His Arg	Val Tyr Ser Phe	Gly
	20	25	30
Gly Tyr Cys Ser Gly	Glu Asp Tyr Glu Thr	Leu Arg Gln Ile	Asp
	35	40	45
Val His Ile Phe Asn	Ala Val Ser Leu Arg	Trp Thr Lys Leu	Pro
	50	55	60
Pro Val Lys Ser Ala	Ile Arg Gly Gln Ala	Pro Val Val Pro	Tyr
	65	70	75
Met Arg Tyr Gly His	Ser Thr Val Leu Ile	Asp Asp Thr Val	Leu
	80	85	90
Leu Trp Gly Gly Arg	Asn Asp Thr Glu Gly	Ala Cys Asn Val	Leu
	95	100	105
Tyr Ala Phe Asp Val	Asn Thr His Lys Trp	Phe Thr Pro Arg	Val
	110	115	120
Ser Gly Thr Val Pro	Gly Ala Arg Asp Gly	His Ser Ala Cys	Val
	125	130	135
Leu Gly Lys Ile Met	Tyr Ile Phe Gly Gly	Tyr Glu Gln Gln	Ala
	140	145	150
Asp Cys Phe Ser Asn	Asp Ile His Lys Leu	Asp Thr Ser Thr	Met
	155	160	165
Thr Trp Thr Leu Ile	Cys Thr Lys Gly Ser	Pro Ala Arg Trp	Arg
	170	175	180
Asp Phe His Ser Ala	Thr Met Leu Gly Ser	His Met Tyr Val	Phe
	185	190	195
Gly Gly Arg Ala Asp	Arg Phe Gly Pro Phe	His Ser Asn Asn	Glu
	200	205	210
Ile Tyr Cys Asn Arg	Ile Arg Val Phe Asp	Thr Arg Thr Glu	Ala
	215	220	225
Trp Leu Asp Cys Pro	Pro Thr Pro Val Leu	Pro Glu Gly Arg	Arg
	230	235	240
Ser His Ser Ala Phe	Gly Tyr Asn Gly Glu	Leu Tyr Ile Phe	Gly
	245	250	255
Gly Tyr Asn Ala Arg	Leu Asn Arg His Phe	His Asp Leu Trp	Lys
	260	265	270
Phe Asn Pro Val Ser	Phe Thr Trp Lys Lys	Ile Glu Pro Lys	Gly
	275	280	285
Lys Gly Pro Cys Pro	Arg Arg Arg Gln Cys	Cys Cys Ile Val	Gly
	290	295	300
Asp Lys Ile Val Leu	Phe Gly Gly Thr Ser	Pro Ser Pro Glu	Glu
	305	310	315
Gly Leu Gly Asp Glu	Phe Asp Leu Ile Asp	His Ser Asp Leu	His
	320	325	330
Ile Leu Asp Phe Ser	Pro Ser Leu Lys Thr	Leu Cys Lys Leu	Ala
	335	340	345
Val Ile Gln Tyr Asn	Leu Asp Gln Ser Cys	Leu Pro His Asp	Ile
	350	355	360

WO 00/78953

PCT/US00/16668

Arg Trp Glu Leu Asn Ala Met Thr Thr Asn Ser Asn Ile Ser Arg  
 365 370 375  
 Pro Ile Val Ser Ser His Gly  
 380

<210> 36  
 <211> 287  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2676443CD1

<400> 36  
 Met Ala Ala Glu Ala Arg Val Ser Arg Trp Tyr Phe Gly Gly Leu  
 1 5 10 15  
 Ala Ser Cys Gly Ala Ala Cys Cys Thr His Pro Leu Asp Leu Leu  
 20 25 30  
 Lys Val His Leu Gln Thr Gln Gln Glu Val Lys Leu Arg Met Thr  
 35 40 45  
 Gly Met Ala Leu Arg Val Val Arg Thr Asp Gly Ile Leu Ala Leu  
 50 55 60  
 Tyr Ser Gly Leu Ser Ala Ser Leu Cys Arg Gln Met Thr Tyr Ser  
 65 70 75  
 Leu Thr Arg Phe Ala Ile Tyr Glu Thr Val Arg Asp Arg Val Ala  
 80 85 90  
 Lys Gly Ser Gln Gly Pro Leu Pro Phe His Glu Lys Val Leu Leu  
 95 100 105  
 Gly Ser Val Ser Gly Leu Ala Gly Gly Phe Val Gly Thr Pro Ala  
 110 115 120  
 Asp Leu Val Asn Val Arg Met Gln Asn Asp Val Lys Leu Pro Gln  
 125 130 135  
 Gly Gln Arg Arg Asn Tyr Ala His Ala Leu Asp Gly Leu Tyr Arg  
 140 145 150  
 Val Ala Arg Glu Gly Leu Arg Arg Leu Phe Ser Gly Ala Thr  
 155 160 165  
 Met Ala Ser Ser Arg Gly Ala Leu Val Thr Val Gly Gln Leu Ser  
 170 175 180  
 Cys Tyr Asp Gln Ala Lys Gln Leu Val Leu Ser Thr Gly Tyr Leu  
 185 190 195  
 Ser Asp Asn Ile Phe Thr His Phe Val Ala Ser Phe Ile Ala Gly  
 200 205 210  
 Gly Cys Ala Thr Phe Leu Cys Gln Pro Leu Asp Val Leu Lys Thr  
 215 220 225  
 Arg Leu Met Asn Ser Lys Gly Glu Tyr Gln Gly Val Phe His Cys  
 230 235 240  
 Ala Val Glu Thr Ala Lys Leu Gly Pro Leu Ala Phe Tyr Lys Gly  
 245 250 255  
 Leu Val Pro Ala Gly Ile Arg Leu Ile Pro His Thr Val Leu Thr  
 260 265 270  
 Phe Val Phe Leu Glu Gln Leu Arg Lys Asn Phe Gly Ile Lys Val  
 275 280 285  
 Pro Ser

<210> 37  
 <211> 497  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3295764CD1

<400> 37  
 Met Asp Val Pro Gly Pro Val Ser Arg Arg Ala Ala Ala Ala Ala  
 1 5 10 15  
 Ala Thr Val Leu Leu Arg Thr Ala Arg Val Arg Arg Glu Cys Trp  
 20 25 30







WO 00/78953

PCT/US00/16668

Thr	Glu	Phe	Asn	155	Trp	Lys	Asn	Val	160	Phe	Ile	Leu	Gln	165
			Gln	170					175					180
Leu	Leu	Ser	Cys	185	Leu	Gly	Phe	Leu	190	Met	Tyr	Ser	Thr	195
			Tyr	200	Asn	Ser	Ala	Leu	205	Thr	Ala	Val	Val	210
Ala	Ile	Lys	Asn	215	Val	Ser	Val	Ala	220	Ile	Gly	Ile	Leu	225
			Phe	230	Ser	Leu	Leu	Asn	235	Val	Gly	Leu	Asn	240
Gly	Asp	Tyr	Ile	245	Leu	Arg	Tyr	Ser	250	Phe	Leu	Thr	Leu	255
			Gly	260	Pro	Val	Gly	Glu	265	Asn	Ile	Cys	Leu	270
Cys	Met	Ala	Gly											
Gln	Leu	Lys	Pro											
Leu	Lys	Ser												

<210> 40  
 <211> 206  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4378816CD1

Met	Gly	Ala	Glu	Trp	Glu	Leu	Gly	Ala	Glu	Ala	Gly	Gly	Ser	Leu
1				5					10					15
Leu	Leu	Cys	Ala	20	Leu	Leu	Ala	Ala	25	Gly	Cys	Ala	Leu	30
				35	Gln	Gly	Ala	Ala	40	Arg	Gly	Ala	Leu	45
Arg	Leu	Gly	Arg	50	Ala	Leu	Val	His	55	Phe	Ala	Leu	Glu	60
			Tyr	65	Ser	Leu	Val	Gly	70	Ala	Asn	Ser	Asp	75
Trp	Leu	Cys	Tyr	80	Leu	Val	Gly	Asn	85	Val	Ala	Asn	Ser	90
			Phe	95	Pro	Thr	Ile	Val	100	Ser	Val	Glu	Ile	105
Phe	Val	Tyr	Leu	110	Gly	Ser	Leu	Ala	115	Phe	Leu	Ile	Tyr	120
			Asp	125	Tyr	Arg	His	Phe	130	Gln	Ile	Thr	Leu	135
Val	Ala	Leu	Asp	140	Gly	Ser	Leu	Ala	145	Thr	Phe	Leu	Pro	150
			Gly	155	Tyr	Gly	Cys	Trp	160	Met	Thr	Phe	Leu	165
Val	Lys	Glu	Lys	170	Trp	Met	Thr	Val	175	Trp	Val	Leu	Ile	180
			Tyr	185	Gln	Ser	Trp	Leu	190	Glu	Leu	Lys	Lys	195
Val	Cys	Glu	Leu	200	Ser	Ser	Val	Lys	205	Phe	Gln			
			Thr											

<210> 41  
 <211> 235  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4797137CD1

Met	Gln	Gln	Arg	Gly	Ala	Ala	Gly	Ser	Arg	Gly	Cys	Ala	Leu	Phe
1				5					10					15
Pro	Leu	Leu	Gly	Val	Leu	Phe	Phe	Gln	Gly	Val	Tyr	Ile	Val	Phe





WO 00/78953

PCT/US00/16668

agagattttc	gatgcactgt	ggggaaactc	tatccatcct	gccttgaaga	gtctccatgg	2280
aaaccaccaa	cttatctttt	ttcaacggat	gggacagaag	agtttgaact	ggatggagaa	2340
atgggtgcac	taccacctgt	atagtgggga	agttcagggg	gtgcacgcct	gagttatgtg	2400
ctttgtcatt	ttctttgcta	aacaaaagag	gctatgaaag	aactaaatat	gagtacataa	2460
aattctatct	ttgataaatt	ttatttttat	gccctactta	atatttgcac	cagtataata	2520
tatatcagtg	agtcttacag	aaagatatgc	ttccataata	tgaaatagat	tattcaataa	2580
ttgagaaact	ttatgtgtaa	tcattgagagt	ataagaatct	ggattatcta	acattgttag	2640
ccctgtgtat	gtacagttca	aaaagtccat	ttataaaaagt	agtttcctgt	tcctagttga	2700
t						2701

<210> 45  
 <211> 736  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1455669CB1

<400> 45						
gagacttagc	gacagacaga	cgctgggacc	cacgacgaca	gaaggcgccg	atggccgcgc	60
ctgctgagcc	ctgcgcgggg	caggggggtct	ggaaccagac	agagcctgaa	cctgccgccca	120
ccagcctgct	gagcctgtgc	ttcctgagaa	cagcaggggt	ctgggtaccc	cccatgtacc	180
tctgggtcct	tggtcccac	tacctcctct	tcattccacca	ccatggcccg	ggctacctcc	240
ggatgtcccc	actcttcaaa	gccaaagatgg	tagctgccat	ccctgggagc	ctggaaccag	300
gcaatgttcg	ggggaggcag	gggacaggct	ggaacctggt	gaagtcttaa	agtagactcc	360
tcctatcggg	gtgtagaagg	gaatctgtta	atcaaacaga	gcaatattag	aaaggctaca	420
gaggtcaact	cagtggaaaca	cggttctccc	aaacagattt	tgtaattccg	aaaatccacg	480
catgcgcaaa	catacgcata	cactcccatg	ttcctggaca	gtttatagct	accataacct	540
ggcattttcc	aaaacatacc	atgtagactc	ttggatacac	aaggtaattt	tagagccaca	600
ttaggatgaa	ccttttaaaa	agttatgcat	ttatttttat	gttccccccac	tggtctgtatt	660
ataggacaat	ttttatatgt	gatatgtatt	taccttagtg	tgttaaataa	acactggcat	720
tcgaagtgtg	aaaaaa					736

<210> 46  
 <211> 1826  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2084989CB1

<400> 46						
ccttaggcgc	cagggacagc	cgagcgttac	ctgggtcccg	gcagcggagt	tctttaccca	60
ccccagttct	ggttctgacg	ccctagctca	ttccgcaaat	ttagggcttg	ggctctggctt	120
gttcccctcc	ggctcgaaac	acctcttctc	tgagccgagc	cagctaccgg	ggctcctgga	180
attgccaccc	ctccctgggc	accttgagg	cctccgtgga	gggacgtcac	ggggcagagc	240
gggacgtgag	cctgagtttg	ctgcaggcgt	gctctgtgtg	gtggctgggt	tctgccaatc	300
cccgtgcccc	ccgggtgggc	gcggccggga	agctcctgcc	cctccctgct	ggctcggcgtc	360
acgcgtgacg	tcccgcgtga	tggtctgggag	ggcccggcgg	cgacagcgga	ggcagagagg	420
aaggcggttc	tgagagcttc	agagagcgat	ggaaagcaaa	atgggtgaat	tgcttttaga	480
catcaacatc	caggaacctc	gctgggacca	aagtactttc	ctgggcagag	cccggcactt	540
tttcaactgt	actgactctc	gaaatctgct	gctgtccggg	gcacagctgg	aagcttctcg	600
gaacatcgtg	cagaactaca	gggcccggcgt	ggtgacccca	gggatcaccg	aggaccagct	660
gtggaggggc	aagtatgtgt	atgactccgc	cttccatccg	gacacagggg	agaagggtgt	720
cctgattggc	cgcattgtcag	cccagggtgcc	catgaacatg	accatcactg	gctgcactgc	780
cacattctac	aggaagaccc	caaccgtggt	gttctggcag	tggttgaatc	agtccttcaa	840
tgccattggt	aactactcca	accgcagtgg	tgacactccc	atcactgtga	ggcagctggg	900
gacagcctat	gtgagtgcc	ccactggagc	tgtggccacg	gccctggggac	tcaaatccct	960
caccaagcac	ctgccccctt	tggtcggcag	atttctgccc	tttgacagcag	tgccagctgc	1020
caactgcac	aacatccccc	tgatgaggca	gagagagctg	caggtgggca	ttccgggtggc	1080
tgatgaggca	ggtcagaggc	ttggctactc	ggtgactgca	gccaagcagg	gaatcttcca	1140
ggtggtgatt	tcaagaatct	gcatggcgat	tcctgccatg	gccatcccac	cactgatcat	1200
ggacactctg	gagaagaaaag	acttccctgaa	ggtaggcgac	tgtaacctctc	ttgtcctgga	1260
atggggcgatg	gctggggagaa	gtgaccaggc	cccaactctc	tctccagcct	cgctgatctc	1320
tctaagactt	gccagccctt	ctcctgaccc	ctgcaccgcc	tcctccacct	tcgttcattc	1380
agcaagaatg	aactggggctg	gggtgaagga	actctgcagg	ggcaggagga	gaggacaaag	1440



WO 00/78953

PCT/US00/16668

```

agggtgatgg agcgccacat ctccctgagat tatttgtacg aattggagca atgttggcctt 1020
atacacctct ggatgagaag agccttgctt tattactcaa ttatcttcac gatttcctaa 1080
agtacctggc aaagaattct gcaactttgt tcagtgccag cgattatgaa gtggctcctc 1140
ctgagtagca tcggaaagct gtgtgagagg cactctcact cacttatgtt tggatctccg 1200
taaacacatt ttgttcttta gtctatctct tgtacaaacg atgtgctttg aagatgttag 1260
tgtataacaa ttgatgtttg ttttctgttt gatttttaaac agagaaaaaa taaaaggggg 1320
taatagctcc tttttctctc tttctttttt tttttcattt caaaattgct gccagtgttt 1380
tcaatgatgg acaacagagg gatatgctgt agagtgtttt attgcctagt tgacaaagct 1440
gcttttgaat gctgggtggt ctattccttt gacactacgc actttttataa tacatgttaa 1500
tgctatatga caaaatgctc tgattcctag tgccaaaggt tcaattcagt gtatataact 1560
gaacacactc atccatttgt gcttttgttt ttttttatgg tgcttaaagt aaagagccca 1620
tcctttgcaa gtcattccatg ttgttactta ggcattttat cttggctcaa attgttgaag 1680
aatgggtggc ttgttctatg tttttgtatt ttgtctaat gcacgtttta acatgataga 1740
cgcaatgcat tgtgtagcta gttttctgga aaagtcaatc ttttaggaat tgtttttcag 1800
atcttcaata aattttttct ttaaatttca aa 1832

```

<210> 49  
 <211> 1211  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 4833111CB1

```

<400> 49
gacagacccg cctcaaacat ggcggcgccc agcgcgcgag gacgtgatcc gcttctgctc 60
cggcttggat tgtagccttg acgaggtctg agcgaccatg gaccggccgg ggctcgtggc 120
agcgctgggtg gctgggtggg tagcaggtgt ttctgttgac ttgatattat ttctctgga 180
taccattaaa accaggctgc agagtcccca aggatttagt aaggctgggtg gttttcatgg 240
aatatatgct ggcgttcctt ctgctgctat tggatccttt cctaattgctg ctgcattttt 300
tatcacctat gaatatgtga agtgggtttt gcatgctgat tcatcttcgt atttgacacc 360
tatgaaacat atgttggctg cctctgctgg agaagtgggt gcctgcctga ttcgagtcc 420
atctgaagtg gttaagcaga gggcacaggt atctgcttct acaagaacat ttcagatttt 480
ctctaacatc ttatatgaag aggggtatcca aggggtgtat cgaggctata aaagcacagt 540
tttaagagag attccttttt ctttgggtcca gtttccctta tgggagtcct taaaagccct 600
ctggctcctgg aggcaggatc atgtgggtgga ttcttggcag tcagcagctc gtggagcttt 660
tgcaggtgga tttgccgctg cagtcaccac ccctctagac gtggcaaaga caagaattac 720
gctggcaaaag gctgggtcca gcaactgctga tgggaatgtg ctctctgtcc tgcattgggt 780
ctggcggtca caggggctgg caggattatt tgcaggtgtc ttccctcgaa tggcagccat 840
cagtcctggga gggttccatc ttctgggggc ttatgaccga acgcacagct tgcgtgttga 900
agttggcaga aagagtcctt gaagcagaga caagcctcac ctccacttct gtcaagagag 960
gggcctgcag tgcaaacctt ctcccgctga gcagctgtct gaactatagg cccagtgct 1020
gaagaccagt tgtgctaaga taccggcatg gagattgtgc catccgtggg ataggctggc 1080
tggtagaag gtattggcct gtatggcaga gagctaagag aagaaaacgg ggtctgtggc 1140
ggtactctga acaatttcct cagaacctct taataaataa gtttggtaat gctgagaaaa 1200
aaaaaaaaa a 1211

```

<210> 50  
 <211> 1046  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 876677CB1

```

<400> 50
cccacgcgtc cggaatgtc tttcactctc tcatactttc tctctctccc tctcccaagc 60
acatctgagt tgctgcctgt tcttcacact tagctccaaa cccatgaaaa attgccaagt 120
ataaaaagctt ctcaagaatg agatggattc taggggtgtc tcacctgaga agcaagataa 180
agagaatttc gtgggtgtca acaataaacg gcttgggtgta tgtggctgga tctgtttttc 240
cctctctttc ctgttgggtga tcattacctt ccccatctcc atatggatgt gcttgaagat 300
cattaaggag tatgaacgtg ctgttgtatt ccgtctggga cgcattccaag ctgacaaaagc 360
caaggggcca ggtttgatcc tggctctgcc atgcatagat gtgtttgtca aagtgaacct 420
ccgaacagtt acttgcaaca ttctccaca agagatcctc accagagact ccgtaactac 480
tcaggttagt ggagttgtct attacagaat ctatagtgtc gtctcagcag tggctaattg 540
caacgatgtc catcaagcaa catttctgct ggctcaaacc actctgagaa atgtcttagg 600

```





WO 00/78953

PCT/US00/16668

ccggaagagc	ccgaagaaga	agatgaagat	gttcaagctg	aaagagtcca	agcagcaa	660
gcactcactg	ctccaaactt	ggaggaggaa	ccagtcataa	ctgcaagctg	tttacacaag	720
gaatatttatg	agacaaagaa	aagttgcttt	tcaacaagaa	agaagaaaat	agccatcaga	780
aatgtttcct	tttgtgttaa	aaaaggtgaa	gttttgggat	tactaggaca	caatggagct	840
ggtaaaagta	cttccattaa	aatgataact	gggtgcacaa	agccaactgc	aggagtgggtg	900
gtgttacaag	gcagcagagc	atcagtaagg	caacagcatg	acaacagcct	caagttcttg	960
gggtactgcc	ctcaggagaa	ctcactgtgg	cccaagctta	ccatgaaaga	gcacttggag	1020
ttgtatgcag	ctgtggaaag	actggggcaa	aaaagatgct	gctctcagta	tttcacgatt	1080
gggtgggaggt	cttaagctcc	aggaacaact				1110

<210> 53  
 <211> 1120  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3735780CB1

<400> 53						
gaaatccagt	tatcaaaatt	gactcaagaa	gagagaacct	aacagaacaa	taacaatgga	60
agaaattggg	aacattatca	caaagctatc	atcctgccaa	actccaggct	cagatgtcac	120
aggttaaaaa	aagtccttca	tgaaaaagaa	agatcttaag	cagcatgatg	gattcagaag	180
ctcatgaaa	gaggcacca	atactaact	cttcaaaaca	agatatatca	cctcatatta	240
caaatgttgg	tgagatgaag	cattacttgt	gtggctgctg	tgagccttc	aacaatgtcg	300
caatcacatt	tcccatccag	aaggctctct	ttcgacaaca	gctgtatggc	atcaaaacc	360
gggatgcaat	acttcagttg	agaagggatg	gatttcgaaa	tttgtatcgt	ggaatccttc	420
ccccattgat	gcagaagaca	actacgcttg	cacttatgtt	tggtctgtat	gaggatttat	480
cctgccttct	ccacaagcat	gtcagtgctc	cagagtttgc	aaccagtggc	gtggcgag	540
tgcttgcaag	gacaacagaa	gcaattttca	ctccactgga	aagagttcag	acattgcttc	600
aagaccacaa	acatcatgac	aaattttacca	acacttacca	ggctttcaag	gcactgaaat	660
gtcatgggat	tggagagtat	tatcgaggct	tggtgccc	tcttttccgg	aatggactca	720
gcaatgtctt	gtttttcggc	cttcgaggct	gcattaaagga	gcactctgcct	accgcaacga	780
ctcacagtgc	tcatctgggc	aatgatttta	tctgtggagg	tctattgggt	gccatgttgg	840
gattcttgtt	ttttccaatt	aatgtttgtaa	aaactcgcat	acagtctcag	attggtgggg	900
aatcttcagtc	ttttcccaag	gttttccaaa	aaactctggc	ggaacgggac	agaaaactga	960
taaattcttt	cagaggtgcc	catctgaatt	acatcggtc	cctcatctct	tggggcataa	1020
tcaatgcaac	ttatgagttc	ttgttaaagg	ttatatgaaa	aaaccatcag	ttaagtgcc	1080
tttatcaact	gaatagacct	tctaagaaga	aaaaaaaa			1120

<210> 54  
 <211> 886  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 039026CB1

<400> 54						
ggccgaggct	cctgtccaga	ccctgaccct	ccctcccaag	gctcaaccgt	cccccaacaa	60
ccgccagcct	tgtactgatg	tcggatgcga	gagcctgtgc	ttaagtaaga	atcaggcctt	120
attggagaca	ttcaagcaaa	ggttggacaa	ctacttttcc	agaacagaaa	ggaaactcat	180
gcatcagaaa	aggtgactaa	taaaggtacc	agaagaatat	ggctgcacaa	ataccagaat	240
ctgatcagat	aaaacagttt	aaggaatttc	tggggacct	caataaaact	acagagacct	300
gctttttgga	ctgtgtttaa	gacttcacaa	caagagaagt	aaaacctgaa	gagaccacct	360
gttcagaaca	ttgcttacag	aaatatttaa	aaatgacaca	aagaatatcc	atgagatttc	420
aggaatatca	tattcagcag	aatgaagccc	tgccagccaa	agcaggactc	cttggccaac	480
cacgatagag	aagtcctgat	ggatgaactt	ttgatgaaag	attgccaaaca	gctgctttat	540
tggaatgag	gactcatctg	atagaatccc	ctgaaagcag	tagccaccat	gttcaacct	600
ctgtcatgac	tgtttggcaa	atggaaaccg	ctggagaaac	aaaattgcta	tttaccagga	660
ataatcacaa	tagaaggtct	tattgttcag	tgaaataata	agatgcaaca	tttgttgagg	720
ccttatgatt	cagcagcttg	gtcacttgat	tagaaaaata	aaccattgtt	tcttcaattg	780
tgactgttaa	ttttaaagca	acttatgtgt	tcgatcatgt	atgagataga	aaaattttta	840
ttactcaaag	taaaataaat	ggaaatatca	ctgaaaaaaa	aaaaaa		886

<210> 55  
 <211> 2336

WO 00/78953

PCT/US00/16668

<212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> misc\_feature  
 <223> Incyte ID No: 260607CB1

<400> 55  
 taatacgcctc actatagggga atttggccct cgagcagtaa ttcggcacga ggaccatctc 60  
 tttaggatat atttttaaat tctttgaaac acataaccaaa aatggtttga ttcactgact 120  
 gactttgaag ctgcatctgc cagttacacc ccaaattggct ttaatcccc ctcgggtctg 180  
 gttgcctttt gcagtttggg ttgtggactc agctcctgtg aggggtctgg ttaggagaga 240  
 gccattttta aggacagggg gttttatagc ccttttctac ttctctcccc tcctccagct 300  
 ccttatcaat cttttttcct ttttcctgac cccctccttc tggaggcagt tgggagctat 360  
 ccttgtttat gcctcactat tggcagaaaa gaccccatTT aaaaccaga gaacactgga 420  
 ggggggatgct ctagtgtggt ctgtgtccat ttctctctgt gccaaagaca gacagacaga 480  
 ggctgagaga ggctgttcct gaatcaaagc aatagccagc ttctgacaca tacctggctg 540  
 tctgaggagg aaggcctcct ggaaactggg agctaagggc gaggcccttc ccttcagagg 600  
 ctcttggggg attagggtgt ggtgtttgct aagccaaggg gtagggagcc gagaaattgg 660  
 tctgtcggct cctggttgca ctttggggaa ggaggggaag tttggggctc cagggtctc 720  
 cctgttggtg gactgctctg tccctgccc ctactgcaga gatagcactg ccgagttccc 780  
 ttcaggcctg gcagacgggc agtgaggagg ggcctcagtt agctctcaag ggtgccttcc 840  
 cctcctccca acccagacat accctctgcc aaactgggaa ccagcagtgc tagtaactac 900  
 ctacagagc cccagagggc ctgcttgagc cttctgtctc cacaggagaa gctggtgct 960  
 ctaggcaacc ccttcctccc acctctcatc aggggtgggg gttctccttt ctttccccctg 1020  
 aagtgtttat ggggagatcc tagtggttt gccattcaaa ccactcgact gtttgctgt 1080  
 ttcttgaaaa ccagtagaag ggaaacagca cagcctgtca cagtaattgc aggaagattg 1140  
 aagaaaaatc ctcatcaatg ccaggggaca taaaagccat ttcccttcca aatactcgac 1200  
 aatttagatg cagaacattt ctctgtattc agacttagag taacaccagc tgaaaactgc 1260  
 agtttctttc ctttggatac ataaggcttc tctatcgggg tacgggacag ggaggaggcc 1320  
 tcatgtctga aggggattt aggggcgaga gccgcagccc tgaccctcgg cctgtgtcac 1380  
 cgctttgggg cacagtctga tggcgcttt gctggcgctc tagtatggtt gactccggat 1440  
 ggacaaaaga aaaaaaattt tttttcttga atgaaatagc aggaagctcc tcgggagcat 1500  
 gtgttttgat taaccgtagt gatggatgct acgagtataa atggattaac tacctcaatc 1560  
 cttacagtaa gattggaact aagggcaggg actcatgcat aagggtatga atcccagcca 1620  
 ggacaagtga gttgaggctt gtgccacaaa aggtttgtcc ttggggaaca ggcaggcctg 1680  
 ccaggatccc ccccatatcg attgggctgg gagggtctggc cgtgaggtcc ccactttctg 1740  
 ctttcttgc ccattgtgta cccctttggc ctccagcttg tccctctctc actttctata 1800  
 gctttgttgg accagatggt gaggaaggga atggcctctt cctttctaga ggggctggc 1860  
 tggagtgaga cctggggctt ggccctggaac ccaccacaca gccccaaagt cagggaagcct 1920  
 ggggaaacca gagctgagac ctcttcaaca ggggttcttt gagatcctac acctccattg 1980  
 ggcccttttt cagtcttcaa tgggggcccc gttggctcta gaaggagaag aggtgaagca 2040  
 ggatcccttt cctgggggga gtctgagggc gcggtccttg gactcattca ggcgtcttt 2100  
 gtagttgggg gagttccact gggcgatccc agccccctcc caccaccct ctaatggacc 2160  
 tcctcataga agccccattt cacttttgtt ttatctacct cttagcaaaa caatagataa 2220  
 attaggtagt ggcagctcca cttgcttagg ttaggggggg aaaaagattt ctttttccaa 2280  
 aggaaaaaaa tattaccttg agaatacttt ccaaaaaata aaatttaaaa aaaaaa 2336

<210> 56  
 <211> 2200  
 <212> DNA  
 <213> Homo sapiens  
 <220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1429651CB1

<400> 56  
 gagaacttta cgcctggatc tcatctaact gacacagaaa ccctgtaagg atccagaggt 60  
 ctcgttcagg accatggaga gcggcaccag cagccctcag cctccacagt tagatcccc 120  
 ggatgcgttt cccagaaagg gcttggagcc tggggacatc gcggtgctag ttctgtactt 180  
 cctctttgtc ctggtgttg gactatggct cacagtgaag accaaaagag acacagtga 240  
 aggctacttc ctggctggag gggacatggt ggggtggcca gtgggtgcat cctgtttg 300  
 cagcaatgtt ggaagtggac atttcattgg cctggcaggg tcaggtgctg ctacgggcat 360  
 ttctgatatca gcttatgaac ttaatggctt gttttctgtg ctgatgttgg cctggatctt 420  
 cctaccatc tacattgctg gtcaggtcac cagcatgcca gaatacctac ggaagcgctt 480  
 cggtggcatc agaatcccc tcatctggc tgtactctac ctatttatct acatctcac 540  
 caagatctcg gtagacatgt atgcaggtgc catcttcac cagcagtctt tgcacctgga 600



WO 00/78953

PCT/US00/16668

```

atagtcacaa aattttacca aaacattaga aacaaaaaat aaggagagcc aagtcaggaa 1860
taaaagtgc tctgtatgct aacgccacat tagaacttgg ttctctcacc aagctgtaat 1920
gtgatttttt tttctactct gaattggaaa tatgtatgaa tatacagaga agtgcttaca 1980
actaattttt atttacttgt cacatttttg caataaatcc ctcttatttc taaattctaa 2040
cttgtttatt tcaaaacttt atataatcac tgttcaaaag gaaatatttt cacctaccag 2100
agtgcttaaa cactggcacc agccaaagaa tgtgggttgta gagaccagga agtcttcaag 2160
aacagccgac aaaaacattc gagttgacct caccaagttg ttgccacaga taatttagat 2220
atttacctgc aggaagggaat aaagcagatg caaccaattc attcagtcga cgagcatgat 2280
gtgagcactg ctttgtgcta gacattgggc ttagcattga aactataaag aggaatcaga 2340
cgcagcaagt gcttctgtgt tctggttagca actcaacact atctgtggag agtaaaactga 2400
agatgtgcag gccaaacattc tggaaatcct atgtcactgg gtttggtttg gaaacctgga 2460
cttctgtcag tttaaaagtt acccagagat gcttctaaag atgagccata gtctagaaga 2520
ttgtcaacca caggagtcca ttgagtggga cagctagata catacattgg cagctacaat 2580
agtatcatga attgcaatga tgtagtgggg tataaaagga aagcgatgga tattgccgga 2640
tgggcatggc cagtgatgtt tcacgtcatt gaggtgacag ctctgtctgga ctttgaatta 2700
catatggagg ctctccagga agacgaagaa gagaaggaca ttctaggcaa aaagaagact 2760
aggcacaagg cacacttatg tttgtctggt agcttttagt tgaaaaagca agatacaggg 2820
tcg                                     2823

```

<210> 58  
 <211> 1491  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2329339CB1

```

<400> 58
cgctccctc cagctgcgag tgcggcctcg gctggcggcg gcaccaggcc acagttgtaa 60
gggatcttgt ggctgtcagg atggcagagg agcaggagtt caccagctc tgcaagttgc 120
ctgcacagcc ctcacaccca cactgcgtga acaacaccta ccgcagcgca cagcactccc 180
aggctctgct ccgagggctg ctggctctcc gggacagcgg aatcctcttc gatgttgtgc 240
tgggtggtgga gggcagacac atcgaggccc atcgcatcct gctggctgcg tcctgcgatt 300
acttcagagg aatgtttgct gggggattga aggaatgga acaggaagag gtctgatcc 360
acggtgtgtc ctacaatgct atgtgccaaa tcctacattt catatacacc tccgagctgg 420
agctcagcct gagcaatgta caagagacac tgggtggctgc ctgccagctt cagatcccag 480
aaattatcca tttctgctgt gatttcctca tgtcctgggt ggacgaagag aacattctcg 540
atgtctaccg gctggcagag ctgtttgact tgagccgct gactgagcaa ctggacacct 600
atatcctcaa aaactttgtg gccttctctc ggactgacaa gtaccgcccag cttccattgg 660
agaaggtcta ctccctcctc agcagcaatc gcctggaggt ctctgcgag accgaggtat 720
atgagggggc ctttctctac cattatagcc tggagcaggt gcaggctgac cagatctcgc 780
tgcacgagcc cccaaagctc cttgagacag tgcggtttcc gctgatggaa gctgaggtcc 840
tgcagcggct gcatgacaag ctggacccca gccctttgag ggacacagtg gccagcggcc 900
tcatgtacca ccggaacgag agcctacagc ccagcctgca gagcccgcaa acggagctgc 960
ggtcggactt ccagtgcgtt gtgggcttcg ggggcattca ctccacgccc tccactgtcc 1020
tcagcgacca gccaagtat ctaaacctct tactgggaga gtggaagcac ttcactgctc 1080
ccctggcccc ccgcatgtcc aaccagggca tcgcggtgct caacaacttc gtatacttga 1140
ttggagggga caacaatgtc caaggatttc gagcagagtc ccgatgctgg aggtatgacc 1200
cacggcacia ccgctgggtc cagatccagt ccctgcagca ggagcacgcc gacctgtccg 1260
tgtgtgttgt aggcaggtac atctacgctg tggcgggccc tgactaccac aatgacctga 1320
atgctgtgga gcgctacgac cctgccacca actcctgggc atacgtggcc ccactcaaga 1380
gggaggtgta tgcccacgca ggcgcgacgc tggaggggaa gatgtatatc acctgcggac 1440
gcaagcttat tccctttagt gagggttaat tttagcttgc actggccgct g 1491

```

<210> 59  
 <211> 986  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2540219CB1

```

<400> 59
ggacgccaac ctcggtttga agtccagggc agtggctcct ggggacagcc agcataccag 60
gggccagtgc actgcattac aaccattgtg aggaatgagg gcctgtgccg tgtggctggc 120
gggcggcatg gcaggagcaa tttcttgggg gacagcgact cctatggatg tcgtgaaaag 180

```



WO 00/78953

PCT/US00/16668

agctcttcct	actgtggaag	gctccagaag	gcccttcaca	aggagacccc	tcacctggat	2640
ccagtcgact	gcggggcctt	cccctcatgt	gggctggcct	ccatcgccca	cgtccaaagc	2700
tgtcactgct	actgcttcag	gctcacatcc	ccccgacctg	atggcggtgc	cgccccctct	2760
ccctgcggcc	catgccacag	gtttctgtgt	tttgcttttag	ggacagaacc	acttaggaaa	2820
gaaagaactc	ccggtctcca	gggtggtatt	tcagtgtctg	tgataatgtc	acgcaacacc	2880
tcttcgggga	ccagtgccca	ggatcctaag	gaagcggaat	tggggcaact	gggcccgtgt	2940
ggccagagct	cagttagcca	gtgccgggcg	gccacagatt	acactgacca	atctcctccc	3000
ttggctctgc	aagcctccca	cccagccttc	tctggcttaa	cccttggttg	cgaaaactct	3060
tccacagtgg	cctccttggg	gacccagaac	ccggaggaag	gggcatgagg	caggaagtgg	3120
ggccgatgtc	tgcaaccag	accacttcgt	ggaatgggct	cttgaccaa	tccctttttt	3180
ttgcgattta	cccgttcaag	caaaacaacg	ttttgggtta	ctaaggattg	tgctaaagcc	3240
gataccaggt	ccttcacacg	tgtgactag	gaacaggagc	gaacagcaca	gagagacgct	3300
ccctgtggga	cgcagcagcc	ccgtggcccc	ggcccagttc	ccagccaccc	tccctggctc	3360
tgctcacacc	agagatttcc	atagcaggag	cggttggtgc	agaagttagt	tcagatgaac	3420
ctcagttaac	gtcgccaccc	ctcctccac	catggtaccc	tgtaggagcc	ctgattgaca	3480
tctgagcgtg	gtggaggtag	gaggggttgc	agctgcagtg	accctgccac	agaggcaggg	3540
tcagtgcaga	ggtcgctttg	gttccgcttc	cctggggcac	agaacggaac	acagcatagg	3600
ttctgcagtc	ggagccgcag	tggcaggatg	gaggggtcga	agggcaagga	gtgactgct	3660
gggcattcct	ggccagcccc	ggccctctgg	tgcctgcttc	ctgtgacttc	agaaggcagg	3720
tggacagagc	ctccctctgg	ccttgctctc	ttcccagcca	cagaacgggc	aggggtggc	3780
ccgacccag	gggagcagta	cctgggtccc	cacccccctc	tcccaaccac	ctccaaggcc	3840
aagctgggtc	ccatagccag	cacggcatgg	ttctccccct	cccccccttc	caggtcaggg	3900
gagttggaca	agtagcaggt	gtttgttttt	aaagcacagc	cctttgggaa	agcaacacat	3960
tattgagact	cactgtgatt	cccccgagg	tcagactggc	tttgtctctt	tctctctgga	4020
ggg						4023

<210> 61  
 <211> 2345  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2739264CB1

<400> 61						
gaaaggaggc	ggaactcggg	gatctgactg	gcgggtttccc	ggccgggactg	agaagggggag	60
cgcgtcgcgc	gtcgaggag	taacctactt	gggtctcctgc	tttcgcgaca	tggccttcaa	120
ttttggggct	ccctcgggca	cctccgggtac	cgtcgcagcc	accgcggccc	ccgcgggttg	180
gtttggaggga	tttggggacaa	catctacaac	tgcaggttct	gcattcagct	tttctgcccc	240
aactaacaca	ggcactactg	gactcttttg	tggtactcag	aacaaagggt	ttggatttgg	300
tactggtttt	ggcacaacaa	cgggaactag	tactggttta	gggtactggt	tgggaactgg	360
actgggattt	ggaggattta	atacacagca	gcagcagcaa	actacattag	gtgggtctct	420
cagtcagcct	acacaagctc	ctacccagtc	caaccagctg	ataaatactg	cgagtgtctt	480
ttctgctcca	acgtgtgttg	gagatgagag	agatgtctatt	ttggcaaaat	ggaatcaact	540
gcaggccctt	tggggaacag	gaaaagggtg	tttcaacaat	aattattccgc	cagtgggaatt	600
cacacaagaa	aatccctttt	gccgatttaa	ggcagtaggt	tatagttgca	tgcccagtaa	660
taaagatgaa	gatgggctag	tggtttttagt	tttcaacaaa	aaagaaaacag	agattcgaag	720
ccaacaacaa	cagtgggtag	aatcattgca	ttaaagtttg	ggaggaaacc	agacccttac	780
tgtaaatgta	gagggcacta	aaacattgcc	agatgacag	acagaagtgt	ttattttatgt	840
tggtgagcgt	tcgccaatag	gtacttcaag	aagagttcca	gctacaacgc	tatatgccc	900
ttttgaacaa	gccaatataa	aaacacaatt	gcagcaactt	gggtgtaaccc	tttctatgac	960
tagaacagaa	ctttctcctg	cacagatcaa	acagctttta	cagaatcctc	ctgctgggtg	1020
tgatcctatt	atctgggaac	aggccaaggt	agataaccct	gattctgaaa	agtttaattcc	1080
tgtagccaat	gtgggtttta	aggaacttct	ccgaagactg	aagggtcaag	atcagatgac	1140
taagcagcat	caaaccagat	tagatatcat	atctgaagat	attagtgagc	tacaaaagaa	1200
tcaaatatac	tctgtagcca	aaattgcaca	atacaagagg	aaactcatgg	atctttccca	1260
tagaacttta	caggtcctaa	tcaaacagga	aagagtgggt	aatgctattca	atgccattca	1320
ggctgatgaa	gagcagttgc	gagttcagct	ggatacgatt	caggggtgaac	taaatgcacc	1380
tactcagttc	aagggccgac	taaatgaatt	gatgtctcaa	atcaggatgc	agaatcattt	1440
tggagcagtc	agatctgaag	aaaggattta	catagatgca	gatctgttac	gagaaatcaa	1500
gcagcatttg	aaacaacaac	aggaaggcct	tagccatttg	attagcatca	ttaaagacga	1560
tctagaagat	ataaagctgg	tcgaacatgg	attgaatgaa	accatccaca	tcagaggtgg	1620
tgctcttagt	tgacagttca	caaacttgtg	taaaggtttg	tgaaatgcat	cttcttactg	1680
catcagacct	tccttaagaa	tgaaccagac	cacatggagg	gaaaaagaaa	acaattcttt	1740
cttggatttg	ttttttgaga	agtttactga	caaattactg	ttcatcaaat	ctgaaatagt	1800
cacctcacag	ctcttcaaag	aaaacctttg	aaagatttat	atctaaaagc	tgtatttact	1860
ttaaaagaag	tgcataatta	ccaaaattgt	atgtactatt	gtacattttt	acaacagcat	1920

WO 00/78953

PCT/US00/16668

tttctttaa	ac	ataatctgtg	tttaatgatt	attgtccatt	gagcctgtac	tctgctttcc	1980
ataccaagta		aatatgaaat	aatctacttt	gcacataaca	gaacaaaacta	taattacttg	2040
gctgttgagg		atttgtactt	gagtataaat	gtacaccagt	ttttgtattt	gtgaactcat	2100
ctgtggggagg		agtaaagaaa	atccaaaagc	atttaaatgtt	ttgtttttgt	tctataaaga	2160
tatgaaaatg		tattttttata	ttatttttact	tatttggaat	ttacagagca	cacctaaagca	2220
attaggatat		aacaaaacta	cttaaccatt	tttgcaacca	ttttgttttt	taagcctttt	2280
tatttctaaa		aagatgaaaa	cttataaata	aattcttaat	ttgtaattac	ttttaaaaaa	2340
aaaaa							2345

<210> 62  
 <211> 2085  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2758310CB1

<400> 62							
cggagatgtg	gcgaccgttt	ctggcatcat	tctgagactc	ggcagttgct	tctcactgct	60	
gcggccgggc	ctgtctgtgg	gagctgcac	ctcctcatct	gcaggcgctg	gaaaaccaga	120	
cacgatcgga	catgcatgtg	gttctgcggc	caaagcacgc	cctttgggtg	tgaacttcat	180	
gatacctgtg	tgcagctctg	tcattttccac	tctgctctgc	tgacacagaag	gcagaagccc	240	
tggccgtccc	ctgctgtgtt	cttcaggaga	aacgtcaggg	gccttctctc	aaggttctcc	300	
agccccacac	ccctgtggag	gaaggtgctc	tccaccgcgg	tagtgggggc	gcccctgctc	360	
ctcggagccc	cttatgtcat	ggcagaggca	cgggagaaga	ggaggatgcg	gctcgtggtg	420	
gatggcatgg	ggcgcttttg	caggctctctg	aaggctcgcc	tgcatatctc	cctggactac	480	
tgggtggtgca	ccaatgttgt	ccttcgaggg	tggagagacc	caggctactt	ggagggtgatg	540	
tctgcgtgtc	accagcgggc	ggctgatgcc	ctgggtggcag	gggccatcag	caacgggggc	600	
ctctacgtga	agctgggcca	ggggctgtgc	tccttcaacc	acctgcttcc	ccccgagtat	660	
acccggaccc	tgcgcgtgct	agaggacagg	gccctcaagc	ggggcttcca	ggagggtgat	720	
gagttgttcc	ttgaggactt	ccaggccctc	ccccacgagc	tcttccagga	gtttgactac	780	
cagccaattg	ctgccgccag	cctggcacag	gtgcacagag	ccaagctgca	cgatggcacc	840	
agcgtggctg	tgaaggtgca	gtacatcgac	ctgcgggacc	gctttgatgg	ggacatccac	900	
accctggagc	tcctgctgcg	gctcgttgag	gtcatgcacc	ccagctttgg	cttcagctgg	960	
gtcctccagg	acctgaaggg	gaccctggcc	caggagctgg	acttcgagaa	tgagggccgc	1020	
aacgcagagc	gctgtgcgcg	ggagctggcg	cacttcccct	acgtcgtggt	gccccgcgtg	1080	
cactggggaca	agtccagcaa	gcgcgtgctc	actgcccact	tctgcgccgg	ctgcaaggtc	1140	
aacgatgtgg	aggccatcag	gagccagggg	ctggcagtg	atgacatagc	agaaaagctc	1200	
atcaaggcct	ttgctgagca	gatattttac	accggcttca	tccactcgga	cccacatcct	1260	
ggcaacgttc	tgggtgcggaa	aggccccggac	gggaaagcgg	agctggtgct	gctggaccac	1320	
gggctctacc	agttcctgga	ggagaaggac	cgcgcagccc	tctgccagct	gtggcggggc	1380	
atcatcctgc	gggacgacgc	cgccatgagg	gcgcacgcag	cgcactggg	ggtgcaagac	1440	
tacctcctgt	tcgccgagat	gctcatgcag	cgccccgtgc	gcctggggca	gctgtggggc	1500	
tcgcacctac	tgagccgcga	agaggcggcc	tacatgggtg	acatggcccc	cgagcgcttc	1560	
gaggcgctca	tggcggtgct	cagggagctg	ccgcggccca	tgtgctggtg	gtcgcgcaac	1620	
atcaacaccg	tgcgcgctat	caacgtggcc	ctcggcgcgc	ccgtggaccg	ctacttcctt	1680	
atggctaaaa	gggctgtccg	gggctggagc	cgccctggcg	gcgccacgta	tcgggggtgtc	1740	
tacggcacca	gcctcctgcg	ccacgcgaag	gtcgtctggg	agatgctcaa	gtttgaagtg	1800	
gcgctcaggg	tggagacctt	ggccatgcgg	ctgaccgcgc	tcctggctcg	tgctctggtc	1860	
cacctgagcc	tcgtgcccc	agcggaggag	ctctaccagt	acctggagac	ctaggggtgca	1920	
gccgcccagg	gccggcgggg	cccttttcac	cttgggctga	cggaggtggc	ggggctagag	1980	
gtgtagacac	cccagacccc	gtgggcactc	gcactggggg	gctgtgacag	cagctggggc	2040	
aggaggccgt	gtaatgacca	cacactcctc	tcaagcaaaa	aaaaa		2085	

<210> 63  
 <211> 3014  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2762348CB1

<400> 63							
ggaggcacaag	gccaggctga	ggatcagggt	ggccccgggtg	gcagcgggga	ggcgctgcat	60	
gctggagggt	gtgctgagtg	cccgggtgcag	gtgagccggg	cctgcggagt	tgtgccgagt	120	
gcctgctgca	gaccgaggct	gggccaagat	ggcgctctgtg	tttcgaagcg	aggagatgtg	180	







WO 00/78953

PCT/US00/16668

<210> 67  
<211> 2574  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 5517972CB1

<400> 67  
gcgcctccca cgccggccgc cgccgacgtg atcgctcggg cgcgccgggc gtgggtgggg 60  
gaaggggttg tgccgcgcga cggctctgct gctgtgcccc ctcaaaaggt tccggggcgc 120  
caggacggaa gacgcagtg cgcacctcc cactgagatt gagagacgcg gcaaggaggc 180  
agcctgtgga ggaactgggt aggatctagg aacgcaccgt gcacatgctt ggtgggtctt 240  
ttaagtggaa actgctgctt tagagtctgt ttggaaggct cgggtgactc atcccaacat 300  
ttacatcctt aattgttaaa gcgctgcctc cgagcgacg catcctgaga tcctgagcct 360  
ttggttaaga ccgagctcta ttaagctgaa aagataaaaa ctctccagat gtcttccagt 420  
aatgtcgaag tttttatccc agtgtcacia ggaacaccca atggcttccc cgcgacagct 480  
tccaatgacc tgaaggcatt tactgaagga gctgtgttaa gttttcataa catctgctat 540  
cgagtaaaac tgaagagtgg ctttctacct tgtcgaaaac cagttgagaa agaaatatta 600  
tcgaatatca atgggatcat gaaacctggg ctcaacgcca tcctgggacc cacagggtgga 660  
ggcaaatctt cgttattaga tgtcttagct gcaaggaaa gttggaactc attatctgga 720  
gatgttctga taaatggagc accgcgacct gccaatctca aatgtaattc aggttacgtg 780  
gtacaagatg atgttgtgat gggcactctg acggtgagag aaaacttaca gttctcagca 840  
gctcttcggc ttgcaacaac tatgacgaat catgaaaaaa acgaacggat taacagggtc 900  
cctcgatatt ccatcttcaa gttgtttgat agctccaagg ttggaactca gtttatccgt 960  
ggtgtgtctg gaggagaaag aaaaaggact agtataggaa tggagcttat cactgatcct 1020  
tccatcttgt tcttggatga gcctacaact ggcttagact caagcacagc aaatgctgtc 1080  
cttttgctcc tgaaaaggat gtctaagcag ggacgaacaa tcattcttctc cattcatcag 1140  
cctcgatatt ccatcttcaa gttgtttgat agctcaccct tattggcctc aggaagactt 1200  
atgttccacg ggcctgctca ggaggccttg ggatactttg aatcagctgg ttatcactgt 1260  
gaggcctata ataaccctgc agacttcttc ttggacatca ttaatggaga ttccactgct 1320  
gtggcattaa acagagaaga agactttaa gccacagaga tcataagacc ttccaagcag 1380  
gataagccac tcatagaaaa attagcggag atttatgtca actcctcctt ctacaaagag 1440  
acaaaagctg aattacatca actttccggg ggtgagaaga agaagaagat cacagtcttc 1500  
aaggagatca gctacaccac ctctctctgt catcaactca gatgggtttc caagcgttca 1560  
ttcaaaaact tgctgggtaa tccccaggcc tctatagctc agatcattgt cacagtcgta 1620  
ctgggactgg ttatagggtc catttacttt gggctaaaaa atgattctac tgggaatccag 1680  
aacagagctg ggggttctct ctctctgacg accaaccagt gtttcagcag tgtttcagcc 1740  
gtggaactct ttgtggtaga gaagaagctc ttcatacatg aatacatcag cggatactac 1800  
agagtgtcat cttatttctc tggaaaactg ttatctgatt tattacccat gaggatgta 1860  
ccaagtatta tattttacctg tacccttatg atggtggctt attcagccag ttccatggca 1980  
gccttctctg cagcaggtca gagtgtgggt tctgtagcaa cacttctcat gaccatctgt 2040  
tttgtgttta tgatgatttt ttcaggtctg cagcattcca cgtatggat ttacggcttt cgagcataat 2100  
ctgtcatggc ttcatgactt ctgcccagga ctcaatgcaa caggaaacaa tccttgtaac 2220  
gaatttttgg gacttgccga agaattttg ggccttggct tgtatgattg ttattttcct cacaattgcc 2340  
tacctgaaat tgttatttct taaaaaatat tcttaaatat ccccttaatt cagtatgatt 2400  
tatcttcaca taaaaaagaa gcactttgat tgaagtattc aatcaagttt tttgtgtgtt 2460  
ttctgttccc ttgccatcac actgttgcac agcagcaatt gttttaaaga gatacatctt 2520  
tagaaatcac aacaaactga attaaacatg aaagaaccca agaaaaaaa aaaa 2574

<210> 68  
<211> 1571  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 5593114CB1

<400> 68  
ggcgcgccga cagctagggt tcacggccac tggggcagag gagccgcgag aagatgtggg 60  
tttttggtta cgggtccctg atctggaagg tggatttccc ctatcaggac aagctggtcg 120  
gatacatcac caactacagc aggcgcttct ggcagggcag cacggaccac cgcgggggtc 180  
ccggcaagcc tgggaagagt gtgactcttg ttgaagatcc tgcgggatgt gtatgggggtg 240

WO 00/78953

PCT/US00/16668

ttgcttacag	attgccagta	ggaaaggaag	agaagtaaa	agcatacctt	gacttcagag	300
aaaaaggagg	ctacagaacc	acaacagtca	ttttttatcc	aaaagatccc	acaacaaaac	360
cattcagtg	attgctatat	attggaacat	gtgataatcc	tgattatctt	ggctcctgcac	420
ctctggaaga	cattgctgaa	caaattttta	atgcagctgg	tccaagtggg	agaaatacag	480
aatatctttt	tgaacttgca	aattctatta	ggaaccttgt	gccagaagaa	gcagatgagc	540
atcttttcgc	tttggaaaaa	ttagtaaagg	aacgttttaga	agggaaacag	aacctcaatt	600
gcatataatt	tagtcttcag	agaattaact	tcagtgcaca	atgacaatat	gatttggaaa	660
tacgtttact	taaagatctt	atttttaatg	tagtgaggat	attattttaa	cttttatttt	720
aactggaaat	gtcctgaaac	acatatttaa	aattattgga	tacagtgaaa	gaaaaattca	780
aatttttaata	acataaagat	ttcctaactt	tatgttattg	aacacttact	cactagaagt	840
gagttcttta	gaaaaataca	gtgaaggact	cagttcagtc	ttgtttttat	cagagtata	900
atcatcctgt	ttcacatccc	aatactattt	tgaaattcta	aacaatttaa	ccaaaattcc	960
aataaatata	aggttatgcc	ttcaatatat	tcctatacaa	ttctgtaacc	atgggtttaa	1020
atacacaagc	ttaaaataac	atgcttagaa	atacacaata	atatgaacag	tatttcagcc	1080
ttaatgtga	atctcctgtg	tattcaagta	ttaaatgaaa	tcttttgagt	ttttagccaa	1140
aaattggcat	ttttaaaata	cgaaaatttc	cttggaatta	taatgtactg	tacctcttg	1200
tttttaaata	aaggcatttt	actatatgga	aaataactca	ctaaagcata	aattacatta	1260
tacaaatcat	gatcactaat	gatgtagtct	gtcattcact	ttgtattaat	cttatacca	1320
aaactgaaaa	gatgggctga	tactacaaat	taatggcaca	tataatgaaa	attagtattt	1380
taaaacagct	tttggaattc	tttgtctgtc	actatctcaa	tttgtgtgtg	tttgtgtgta	1440
tatacataaa	tatacatata	aaattttttt	tttctttgca	gcctgcgtct	ggccatccca	1500
caggctggaa	agtgtaacct	ctggcagaag	ccaagaacag	gcacctcctg	gaattataat	1560
tttgtttgt	t					1571

<210> 69  
 <211> 1549  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 044775CB1

<400> 69						
cgcgcttagg	caggcggtgg	cgcggtcgga	gtgccgcggg	gagggctgtg	ccggttgctt	60
tctgcagccg	catctcggcc	agctctcctc	gccgtccccg	ggcgctgtg	cgtctccagt	120
ccgggaccga	agccgcctgc	cgtagcgggc	ggccagatcc	gcgtcccgc	tcagcggccg	180
gaggacatgc	gggagagaga	atgagccaga	gggacacgct	ggtgcatctg	tttgccggag	240
gatgtggtgg	tacagtggga	gctattctga	catgtccact	ggaagtgtga	aaaacacgac	300
tgcagtcatc	ttctgtgacg	ctttatatatt	ctgaagttca	gctgaacacc	atggctggag	360
ccagtgtcaa	ccgagttagt	tctcccggac	ctcttcattg	cctaaagggtg	atcttggaaa	420
aagaaggggc	tcgttccttg	tttagaggac	taggccccaa	tttagtgggg	gtagccctt	480
ccagagcaat	atactttgct	gcttattcaa	actgcaagga	aaagtgtgaat	gatgtatttg	540
atcctgatcc	tacccaagta	catatgattt	cagctgcaat	ggcaggcatt	tactgtacat	600
ttctcccag	aaaagagtga	gatcgtgtca	tctcatgctc	cccacccgca	ggtcacttcc	660
tgtagaaata	tggactaact	taaacctcgt	tttactgcaa	tcacagcaac	caacccctt	720
tggcttataa	agactcggtt	acagcttgat	gcaagggaacc	gcgggggaaag	gcgaatgggt	780
gcttttgaat	gtgttcgtaa	agtgtatcag	acagatggac	taaaaggatt	ttataggggc	840
atgtctgctt	catatgctgg	tatatcagag	actgttatcc	attttgttat	ttatgaaagt	900
ataaaaacaaa	aactactgga	atataagact	gcttctacaa	tggaaaatga	tgaagagtct	960
gtgaaagaag	catcagattt	tgtgggaatg	atgctagctg	ctgccacctc	aaaaacttgt	1020
gccacaacta	tagcatatcc	acatgaagtt	gtaagaacaa	gactacgtga	agagggaaca	1080
aaatacagat	ctttttttca	gactctatct	ttgcttgctt	aagaagaagg	ttatgggtct	1140
ctttatcgtg	gtctgacaac	tcacttagtg	agacagattc	caaacacagc	cattatgatg	1200
gccacctatg	aattgggtgg	ttacctactc	aatggatagc	agcacgagga	ctgctgtact	1260
gcaaaaaaag	aagaccaaaa	gattacagtg	gaccatggga	tacagaagcc	agcatggcag	1320
acagaagaaa	aatagtttgg	gaacatgtaa	ctatttctaag	tggaaagtgt	gtttagggaa	1380
ttatagtaat	cacaccacat	tacttggcct	ttcggtaatg	tgaaaaaaaa	aaaaaaacct	1440
cagagcctcc	aaggaaatgc	ctttagaagc	actctctctc	caaaattgcc	attttctcta	1500
ccatgtcccc	cagacacagt	tgggttttgt	tgatttatgg	cagtccttct		1549

<210> 70  
 <211> 2237  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature

WO 00/78953

PCT/US00/16668

<223> Incyte ID No: 116588CB1

<400> 70

```

gtcactgggtg ccacgggggcc tcagcgcact tctgtctttaa gctcctgggc ctccttattt 60
tcccccttgc gatcgattcc agccacacct gtggatgttg ctagttactc cgcgtccgga 120
acgtggaggt cgagggactg agctggggcga gttttgtggc actcctttgc tcttcagcag 180
ctatttttgc tatgataatc ctgctgccct tcagactcaa gttaaaccgag atagcaagt 240
gaataccacg aaattcatgc tgctgtatgc ctgggtattct tggcccaatg tagttttgtg 300
tttcttttgg ggttttttga tagaccgagt attttgaata cgatggggca caatcatttt 360
tagctgcttt gtttgcattg gacaggttgt ttttgccttg ggtggaatat ttaatgcttt 420
ttggctgatg gaatttggaa gatttgtatt tgggtttgtt ggcgagtcct tagcagttgc 480
ccagaatata tatgtcttga gctggttttaa aggcaaagaa ttaaaccctg tgtttggact 540
tcaacttagc atggctagaa ttggaagtac agtaaaccat aacctcatgg gatggctgta 600
ttctaagatt gaagctttgt taggtttctgc tgggtcacaca accctcggga tcacacttat 660
gattgggggt gtaacgtgta ttctttcact aatctgtgcc ttggctcttg ctagcttggg 720
tcagagagca gagagaatcc ttcataaaga acaaggaaaa acaggtgaag ttattaaatt 780
aactgatgta aaggacttct ccttaccctt gtggcttata tttatcatct gtgtctgcta 840
ttatgttgcg gtgttccctt ttattggact tgggaaagtt ttctttacag agaaatttgg 900
attttcttcc caggcagcaa gtgcaattaa tggtttgtta tatgtcata tagctccat 960
gtccccgggtg tttgggctcc tgggtggataa aacagggaag aacatcatct ggggttctttg 1020
cgcagtagca gccactcttg tgtcccacat gatgctggcc tttacgatgt ggaacccttg 1080
gattgctatg tgtcttcttg gactctccta ctcatgtctt gcctgtgcat tgtggccaat 1140
gggtggcattt ttagttcctg aacatcagct ggggaactgca tatggcttca tgcagtccat 1200
tcagaatctt gggttggcca tcatttccat cattgtctgt atgatactgg attctcgggg 1260
gtatttgttt ttggaagtgt tcttcattgc ctgtgtttct ttgtcacttt tatctgtggt 1320
cttactctat ttggtgaate gtgccagggg tgggaacctt aattattctg caagacaaag 1380
ggaagaaata aaattttccc atactgaatg agaagttaaa atgaatgtgt catgagaatg 1440
ggcttaacac atcgttgggt tgaaaacttc cattttttaa aatttagagt ttagtcatta 1500
gaaaaaataa tggactggaa agtttatatt atatccaaat atacctattt caaagtgtat 1560
ttgtgagctc tgttttagcc tgtgtctttt gtattgtgtg ttgctaaaga attctacttt 1620
tagtaggcta atcaacaatg aaagggttag ggaagcgtt ggagcttctg ttgagataat cttcattagg 1680
caggaaagac agtgaaaaat ggaaaacgtt atcttatagc agggaaaaaa aactttttgag 1740
tatatatctt agggatacag ccttttcttt acataagga aacaaatggc ttgtcatagg ccttcctttt 1800
ggaaatagaa ggcctgcgtt acacaaaata ggggaatatg gaagatgttt ttatgaagtc tctttctgat 1860
actagtaggg cataatgcta ggggaatatg ctactctgta gttatgtgga ttgccgagca atgacccttt 1920
cacgaacaat agcttgcgct tactgaggac ctaatacact tagggatgta attttatagt 1980
tcaatttctt atttctgtgt tctttagagt tctttagagt aaaagtgtcc ttcaaattat 2040
ataaactttc tgtacagttt ttcttagagt gataaattaa aacactgcac acggagtaaa aaaaaaaa 2100
gataattgcc tatgtacatg gataaattaa aacactgcac acggagtaaa aaaaaaaa 2160
aaaaaaaaaa aaaaaatgag cggccgcaag cttattccct ttagtgaggg ttaattttag 2220
cttggcactg gccgtcg 2237

```

<210> 71

<211> 1114

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 875369CB1

<400> 71

```

tgggaacgga gcagccccgg gggccccctt gagggcgga ggccgcgaag ggcgcggggc 60
tggagggtag gagagcgcg gaaagcgccc cagacgccac tcgcggcgga cggcgggcag 120
ttccagggg tttggagcgg gcccgcggcg ccattggtca cgtcggctcc cgcaagcgt 180
cgaggagtcg cagccggtcc cggggacggg ggtcggaaaa gagaaagaag aagagcagga 240
aagacacctc gaggaactgc tcggcctcca cateccaagg tcgcaaggcc agcacggccc 300
ctggggcgga ggcctcacct tctccctgca tcacagagag aagcaagcag aaggcccgga 360
ggagaacaag atccagctcc tctcctctct ctccagttc ttctagctcc tcttcttctt 420
cctcgtcctc ctccctcttc tccagtgtat gccggaagaa gcgggggaag tacaaggaca 480
agaggaggaa gaagaagaag aagaggaaga agctgaagaa gaagggcaag gagaaggcgg 540
aagcacagca ggtggaggct ctgccgggcc cctcgttgga ccagtggcac cgatcagctg 600
gggaggaaga ggatggccca gtccctgacg atgagcagaa gtcccgaatc caggccatga 660
agcccatgac caaggaggag tgggatgccc ggcagagcat catccgcaag gtgggtggac 720
ctgagacggg gcgcaccagg cttattaagg gagatggcga ggtcctagag gaaatcgtaa 780
ccaaagaacg acacagagag atcaacaagg aagccaccgg aggggactgc ctaggcttcc 840
agatgcgagc tgggttgctt ccctgagggc ccccgctggc caaggcctgt ggacgacgct 900
ggcgggccag cctgggcagg tttcagggtg ccagtgggaa gcctgatggg tgcgtggggc 960

```

WO 00/78953

PCT/US00/16668

ctttcccccg tggattgggc tctggcccag cccagtctct tctcaggggc aggggggtgga 1020  
ggttgggggc accggcctgc ttggcaccac catctgaaag agcagcactt ctcagctatt 1080  
aaaggccccc tggatagaca aaaaaaaaaa aaaa 1114

<210> 72  
<211> 998  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 1325518CB1

<400> 72  
gactaaaaaa gccatgtatt ctttcgtttc tctctaaaag aagaaaaata taatttataa 60  
atacattgcg tattttctaa aacaataaat ttatagtgtt aatattcata ggggtcaatca 120  
aaatgaagct tctccttttg gcctgcattg tatgtgttgc ttttgcaagg aagagacggg 180  
tccccctcat tgggtaggat gacaatgacg atgggtaccc acttcatcca tctctgaata 240  
ttccttatgg catacggaat ttaccacctc ctctttatta tcgcccagtg aatacagtc 300  
ccagttaccc tgggaatact taaactgaca cagggttacc ttctgtatccc tggattctaa 360  
cttctcctgg attccccctat gtctatcaca tccgtgggtt tcccttagct actcagttga 420  
atgttccctc tctccctcct aggggtttcc cgtttgtccc tccttcaagg tttttttcag 480  
cagctgcagc acccgctgcc ccacctattg cagctgagcc tgctgcagct gcacctctta 540  
cagccacacc tgtagcagct gagcctgctg cagggggccc tgttgcagct gagcctgctg 600  
cagaggcacc tgttggagct gagcctgctg cagaggcacc tgttgcagct gagcctgctg 660  
cagaggcacc tgttggagct gagcctgctg cagaggcacc ttcaccagct gagcctgcta 720  
cagccaagcc tgctgcccc gaacctcacc ctctccctc tcttgaacag gcaaatcagt 780  
gaaattctct agaagagtac catgggttca tttctatact gatgcagaaa taagtgaat 840  
ctacaaaagt tttctttctt ttccaaagac tatttcattc tgttgtattc agagtattca 900  
tctcactaca ttgatttgtt tgtggtagtt tttccttggc cttaatttat attgaaaaaa 960  
cattgataat taaataaata aaatagataa ttttagaca 998

<210> 73  
<211> 2348  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<223> Incyte ID No: 2060987CB1

<400> 73  
gggccagccg gctcgccccg gggccatggc agcagcggct actgcagccg aggggggtccc 60  
cagtcggggg cctccccggg aagtcaccca tctgaatgtg ggaggcaaga gattcagtag 120  
ctctcgccag actctcacct ggatcccaga ctctctcttc tccagctctc tgagcgggacg 180  
catctcgacg ctgaaagatg agaccggagc aatcttcatc gacaggggacc ctacagctct 240  
cgcccccatc ctcaacttcc tgcgcaccaa agagttagat cccaggggtg tccacgggtc 300  
cagcctcctc catgaagccc agttctatgg gctcactcct ctggttcgtc gcctgcagct 360  
tcgagaggag ttggatcgat cttcttgtgg aaacgtcctc ttcaatggtt acctgccgcc 420  
accagtgttc ccagtgaagc ggcgggaacc gtggggccct agcagctagg 480  
aggacggcca gcccctgtcc gacggagcaa cacgatgccc cccaaccttg gcaatgcagg 540  
gctgctgggc cgaatgctgg atgagaaaac ccctccctca ccctcaggac aacctgagga 600  
gccgggggat gtgcgcctgg tgtgtggaca ccataattgg atcgctgtgg cctataccca 660  
gtttctagtc tgctacaggt tgaaggaagc ctctggctgg cagctgggtg tttccagccc 720  
ccgcctggac tggcccatcg aacgactggc gctcacagcc cgggtgcatg gtggggcttt 780  
gggtgaacat gacaagatgg tggcagcagc caccggcagc gagatcctgc tatgggctct 840  
gcaggcggaa ggcgggtggc ccgagatagg ggtctttcat ctgggggtgc ctgtggaggg 900  
cttgttcttc ctgcccgaacc agctcattgc tacaagccac acaggggcga tcgggggtgtg 960  
gaatgccgtc accaagcact ggcagggtcca ggaggtgcag cccatcacca gttatgacgc 1020  
ggcagggtcc ttcctcctcc tgggctgcaa caacggctcc atttactacg tggatgtgca 1080  
gaagtcccc ttgcgcctga aagacaacga cctccttgtc agcgagctct atcgggaccc 1140  
agcggaggat ggggtcagcg ccctcagttg ctaccctacc cccaagacca ctgacagtgg 1200  
gaactggatc gagatcgctc atggcaccag ctcagggggc gtgcgggtca tcgtgcagca 1260  
cccggagact gtgggctcgg ggcctcagct cttccagacc ttcactgtgc accgcagccc 1320  
tgtcaccaag atcatgctgt cggagaagca cctcatctca gtctgtgccc acaacaacca 1380  
cgtgcggaca ttgctgtgta ctgccttccg cgccatgatt tccaccagc cgggtccac 1440  
cccactcgct tcccttaaga tcttggctct ggagtcggca gatgggcatg gcggtgcag 1500  
tgctggcaat gacattggcc cctacgggtg gcgggacgac cagcaagtgt tcatccagaa 1560









WO 00/78953

PCT/US00/16668

tctcctggggc	catggcagcc	gagggcgcgcg	tgctcgcgctg	gtacttcggg	gggctggcct	240
cctgcgggggc	cgcttgcctgc	acgcacccgc	tggacctgct	caaggtgcat	ctgcagacgc	300
agcaggaggt	gaagctgcgc	atgacgggca	tggcgctgcg	ggtgggtgcgt	accgacggca	360
tcctggcact	ctacagcggc	ctgagcgccct	cgctgtgcag	acagatgacc	tactccctga	420
ctcggttcgc	catctacgag	actgtgcggg	accgcgtggc	caagggcagc	caggggcctc	480
tccctttcca	cgagaaggtg	ttgctgggct	ccgtcagcgg	tttagctgga	ggcttcgtgg	540
ggacgccccg	agacttggtc	aacgtcagga	tgcagaacga	cgtgaagctg	ccccagggtc	600
agcggcgcaa	ctacgccccat	gcgctggatg	gcctgtaccg	cgtagctcgt	gaagaggggc	660
tcaggagact	gttctcgggt	gcaaccatgg	catccagccg	aggggcctta	gtcactgtgg	720
gccagctgtc	ctgctacgac	caggccaagc	agctggctct	tagcaccggg	tacctctctg	780
acaacatctt	cactcacttt	gtcgccagct	ttattgcagg	tggatgtgcc	acgttcctgt	840
gccagccctt	ggatgtgctg	aagactcgcc	tgatgaactc	caagggggag	tatcagggcg	900
ttttccactg	cgccgtggag	acagcgaagc	tcgggcctct	ggccttttac	aagggcctcg	960
tcccagctgg	catccgcctc	atccccca	ccgtgctcac	ttttgtgttt	ctggaacagc	1020
tacgcaaaaa	ctttggcatc	aaagtggcat	cctgaccagc	cgtgggaatg	gctgggctgc	1080
caggccagac	acgctaggtt	cttccaaaga	gtcccaagcc	cagcacctgc	tcctggggcc	1140
acgacctccc	tggccgtggc	cacccatcct	ccgcagcagg	cccctgctgt	ccccccacct	1200
gctggctgag	ctcctcctgg	cctcgteccc	ttcagctgtt	agctgcacca	cccccgctct	1260
ggctaccagg	ctctcccggc	tgggcactgc	gtggccttgc	ccctctcccc	ctggcagctc	1320
ctcaggggaa	caggggctac	cagaggetga	tttctccctt	ctcctggggc	aggggagggg	1380
tattatccct	gcctcctgcc	cccgatgccc	aaagcagcat	cttcacgac	tttccatcga	1440
ggacttgggt	ggcagagtgt	gggtgcagcc	tggctgttgc	tcacccaagt	gctagctctg	1500
cacttcgtgt	ctgctgagag	caaccagacc	ttccatgtcc	tcgggcagct	gcaacctccc	1560
gcgagacccc	gcagctgggt	gggatgaaca	agcaacgcag	accacaagcg	agtgcctggg	1620
agggagtggc	ccaggggtgg	tctggagcca	ttgtgggtga	gggtcgaggg	ccaccgaggt	1680
cccgcgcacc	gctgcctgcc	ctgcagtggc	tttaacagtt	agttttgcca	aagcctctcc	1740
actcaccagc	agggcgtctc	tgtcttcagg	gattgtgcct	gcgtccctcg	ggcactctgg	1800
cccccccgct	tggctccctg	ggggaatggc	ccaggcgggc	cgcggttctt	ccttagggcc	1860
ttctccccga	caaggagtcc	gacggggcgg	atgctgcac	ctctgcctcc	ctggtcgtctg	1920
ggcttcaccc	cacctgggaa	gggcagtgtg	ctctgtgggg	gctgcaatca	ataaatgccg	1980
ggagctgcca	aaaaaaaaaa	aaaa				2004

<210> 80  
 <211> 3555  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3295764CB1

<400> 80						
gaccgtggggc	gagtcagaa	cgctcctggcc	ttacagggag	aaggcgctcac	tcgcggttac	60
aagtgcctga	ccctcactcc	agttggcgga	ggaggagaag	gaagggggcg	ggccgggtcc	120
cctccccctcg	cgccccggat	ggatgtgccc	ggcccggtgt	ctcggcgggc	ggcgggcgcg	180
gcggccactg	tgctcctgcg	gacgcctcgg	gtccgctcgcg	aatgctgggt	cttgcgcacc	240
gcgctgctct	gcgcctacgg	cttcttcgcc	agcctcaggc	cgtccgagcc	cttcttgacc	300
ccgtacctgc	tggggccgga	caagaacctg	accgagaggg	aggtcttcaa	tgaattttat	360
ccagtatgga	cttactctta	cctggtgcta	ctgtttcctg	tgttccttgc	cacagactac	420
ctccgttata	aacctgttgt	tctactgcag	gggctcagcc	ttattgttac	atggtttatg	480
ctgctctatg	cccaggggact	gctggccatt	caattttctag	aattttttta	tggcatcgcc	540
acagccactg	aaattgccta	ttactcttat	atctacagtg	tgggtggacct	gggcatgtac	600
cagaaagtca	caagttactg	tcgaagtgcc	actttgtgtg	gctttacagt	gggctctgtc	660
ctagggcaaaa	tccttgtctc	agtggcaggc	tggctcgtgt	tcagcctgaa	tgtcatctct	720
cttacctgtg	tttcagtggc	ttttgctgtg	gcctgggttt	tacctatgcc	acagaagagc	780
ctcttctttc	accacattcc	ttctacctgc	cagagagtga	atggcatcaa	ggtacaaaat	840
ggtggcattg	ttactgacac	cccagcttct	aaccaccttc	ctggctggga	ggacattgag	900
tcaaaaatcc	ctctaaaatc	ggaggagcct	cccgctggag	aaccggaacc	cagccagagc	960
cgtctccttg	tattgaaagt	actatggaa	gatttcctga	tgtgctactc	ctctcgccct	1020
cttctctgct	ggtctgtgtg	gtgggcccct	tctacctgtg	gctattttca	agttgtgaac	1080
tacacacagg	gcctgtggga	gaaagtgtgt	ccttctcgtc	atgctgctat	ctataatggt	1140
ggcgtggagg	ccgtttcaac	cttactgggt	tggttgcgtg	tgtttgcagt	tggttatata	1200
aaaatatacct	ggtcaacttg	gggagaaatg	acattatctc	tcttttctct	cctgattgct	1260
gctgcagtgt	atatcatgga	cactgtgggt	aacatttggg	tgtgctatgc	atcctatgtt	1320
gtcttcagaa	tcactacat	gttactcacc	acgatagcaa	cttttcaaat	tgttgcaaac	1380
ctcagctatg	aacgctatgc	cctagtattt	ggtgtaaata	ccttcattgc	ctcggcactg	1440
cagacgctgc	tcactcta	tgtggtagat	gccagtggcc	ttggattaga	aattaccact	1500
cagtttttga	tctatgccag	ttatttttga	ctcatcgctg	tggttttcct	ggccagtggg	1560

WO 00/78953

PCT/US00/16668

accacttcat aatatactgc tgaagggcct cttcttatag caagaactct gcacagcaac 1680  
 tgccctggatg tatttgattt tttaaagcgt agacatatat ttatgaatgt gcatttcttg 1740  
 acttcacagc agccacttga ctaatacctt gtgttccggg aataacatga tactattcag 1800  
 aggagccaga agtaaaagttt atttcatgga ttatttatga gagctaattt aaggatgact 1860  
 ttttttctga ttcaaaagtg aacttgattt taaaaaccag tcaagagcaa tcaaaagcagc 1920  
 acatgggtgtt gtatacttca ttagcaagtgt agtttggtgt tttataggtc acatatgtct 1980  
 gtatctactt agccagatgc ttggcctggg gggaccaggg ctccacagag gccacaaaat 2040  
 gttgcaagtc atgatggatg gaaatatgtt ctaacagcat ctgcctctat tcaatttaat 2100  
 tcttatttct gtgttactca tgtacattgg tctttctaca tagttattct atcactggca 2160  
 atatttggtc ttctgtattt taagggtgtac gtatcatttc taattttaag 2220  
 ttattttaaa aaaattcatc atatgaatgt tcttggttcc cattgtgacg attatttatt 2280  
 tctgtaaaaat ttgttttagaa gtacgttttt gcattattca tatgcttccc agagaagctc 2340  
 atttagttag aaaataaggc aagttttgaa gcctgtctaaa tgaagagact taagaaagct 2400  
 taaggtaacg ttgcttgtct ttaaactctc aatatgaagg actattaatt ccaagattaa 2460  
 aagttcatat ataggctaaa gatgtaacta ggccatttgt atttgtattc ccttttattt 2520  
 ccaaaaataaa atgaaaaatc tttttttaat aatttcatcc ctatttatag tttttatatt 2580  
 aatttggttt tcttatccaa gtaaagatgt caataggaat tgcattagtc caaggccttt 2640  
 ttcataaact gagcctcttt tcaattattt caatgggaca ggaactagga tagatgtgat 2700  
 tcttgcattt ttttacctta aatctgcctt tgtttctaaa ggtagatcat cttgaatatt 2760  
 tgcttaaaat tgctagtgat ttcattacca agttacttga aaaaatgttc tatatgcatt 2820  
 taattctgaa atcagtctac caaggggctg ctagtatatg tcagacatga aaactatttt 2880  
 aaagctgact ttgttgccct atcttgaaaa gaacttagat aggtgctttt aactggggta 2940  
 ttaacttttt tagaatgaca cagctgaaca gtgttaataa tagtgtgtca agattgcaa 3000  
 gtcgacatac tcatttgggt taagcaggaa tcctagaagc aaatggatgg ggataagaat 3060  
 aggtcatttt ctattcacca tcctttacta ttaagggaaa ggaaaagaac actagctaag 3120  
 gaagggaaa ggaagtgata tcataaaaagt agcaaccttc attttacatt ctgtctgttt 3180  
 ttcttttttt gctttgtttt gtttgtgcta atttgggaat tgtgtactcc gaaacaagta 3240  
 gaaaagtgtt gtttgaggga ttttattaaa tcttttttta atggaatgtg gtacaaattg 3300  
 ttcatgttac caaagcaata tttccctgga atttaattca aagtttgtgg catacaacct 3360  
 gagccttttc ttatataaga caagaatatg ttcacatctt ggtatgtggc catatttata 3420  
 gaatgctgaa ctcaatgtgc aagttgtact gtatgcagtt ttgtaataa gtgaaaataa 3480  
 tttgtgtgac tttttattca attctgtata gattataaaa ttatttttat taaataaata 3540  
 ttttacagta tattt 3555

<210> 81

<211> 1293

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<223> Incyte ID No: 3438320CB1

<400> 81

atatgcccc cgcgggactc atactacgtt tcccgtgaac acgtgcagtc caaaccccg 60  
 ccctgatatt tatctcagtg gacggtggcc ggaaaaggac aatgggtttcc atgtcagcgg 120  
 ataaacgctc tcccctcggc tccggagcgc gacggagggtc gtagtagtag tgagtacgtg 180  
 ctgaggagca aaggagtaac caagagatcc agtgaccgac agagcaagag ccatgcccg 240  
 ccggggcctg gtggctgggc cagacttggc cctcttccag cgtcgctatt tcacgcccga 300  
 ggaggtggcc caacataaca ggcccgaaga cctctgggta tcttacctgg gacgcgtgta 360  
 cgacctaacg tcattggcac aggaatacaa ggggaacctg ctgctgaaac ccacgtgga 420  
 agttgcaggc caggatatca gccactgggt tgatccaaag accagagaca tccgcaagca 480  
 catagatccg ctgaccggct gcctgaggta ctgcaccccg cggggcccgt ttgtgcacgt 540  
 tccgcctcag ctgcccgtgt cggactgggc caacgatttt gggaagccct ggtggcagg 600  
 gtcgtattat gaggtggggc ggctgtctgc caagaccccg agcatccgca tcattaacac 660  
 gctcacgtcg caggagcaca cactggagggt ggggggttctg gagtccatat gggaaatcct 720  
 acaccgctg acaccctata actcacatgc tgccagctac acgtggaaat tcacaggga 780  
 gaacctgaac atggatttta ccctggaaga gaatgggatc cgggatgagg aggaagaatt 840  
 tgactatctc agtatggacg gtacacttca cacacctgca atacttctgt acttcaatga 900  
 tgatctcagc gagttgtagg caaggagatg tacactcgtg tagactcaag acgtatttctg 960  
 agtttgctt ttctgtgccc ttgaggaaa gctgtggggc cgaggggtgc cgagggccag 1020  
 atctccactc ctctccagga gctagcctgt gcccttctga agtgtaaagg ccctattccc 1080  
 tgccttcatt acagtttgct ctgagaaaat tagtgaaata atctttggga atgatacaag 1140  
 aagatcaagt accttggttt agggagatgt agaagaggat agtcagagtt caggcagaac 1200  
 tgtttgatag ttaagagaga gtatgtctac aggggtgagg gatggaagga cttttttggc 1260  
 aatgatggaa atgagatgtc tgcaggagat ggg 1293



